



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 155849

TO: Medina A Ibrahim
Location: REM-2B07/2C18
Art Unit: 1638
Thursday, June 09, 2005

Case Serial Number: 10/718952

From: Toby Port

Location: Biotech-Chem Library
REM1-A59

Phone: 272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Ibrahim,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

This Page Blank (uspto)



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or *contact:*

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop-off or send completed forms to STIC Biotech-Chem Library, Remsen Bldg.



This Page blank (uspro)
BEST AVAILABLE COPY

This Page blank (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:31:27 ; Search time 31 Seconds
(without alignments)
1228.098 Million cell updates/sec

Title: US-10-718-952-12
Perfect score: 2631
Sequence: 1 MEIENFKVESPVKXTETETI.....NIMRACVGLAPENNILEYK 510

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2346	89.2	510	3	US-09-118-442-11
2	2346	89.2	510	3	US-09-677-064-11
3	2337	88.8	510	4	US-09-727-628-2
4	1389.5	52.8	533	4	US-09-734-237B-73
5	1389.5	52.8	534	4	US-09-734-237B-75
6	1384.5	52.6	525	4	US-09-248-796A-17234
7	528	20.1	444	4	US-09-902-540-12518
8	119.5	4.5	452	4	US-09-248-796A-20645
9	118.5	4.5	2442	4	US-09-077-098A-6
10	116	4.4	730	4	US-09-107-433-4707
11	116	4.4	731	4	US-09-583-110-4720
12	115	4.4	842	4	US-09-489-039A-11570
13	114	4.3	731	1	US-09-014-897-2
14	114	4.3	731	1	US-08-731-716-2
15	113.5	4.3	778	4	US-09-489-039A-8455
16	110.5	4.2	747	4	US-09-583-110-3066
17	110.5	4.2	751	4	US-09-107-433-3375
18	110.5	4.2	1563	4	US-09-583-110-5243
19	110.5	4.2	1567	4	US-09-107-433-4883
20	110	4.2	935	4	US-09-512-250C-33
21	109.5	4.2	835	4	US-09-438-185A-728
22	109.5	4.1	1242	4	US-09-107-532A-5241
23	105	4.0	2465	2	US-08-596-291-3
24	105	4.0	2465	3	US-09-100-804-3
25	105	4.0	2466	3	US-09-080-855-12
26	105	4.0	2466	4	US-09-566-076-12
27	105	4.0	2466	5	PCT-US94-09943-2

28	105	4.0	2485	3	US-09-290-640-46	Sequence 46, Appl
29	105	4.0	2485	4	US-09-665-615B-46	Sequence 46, Appl
30	104.5	4.0	545	4	US-09-248-796A-14131	Sequence 14131, A
31	102.5	3.9	778	4	US-09-792-024-112	Sequence 112, App
32	101.5	3.9	604	3	US-08-569-749-4	Sequence 4, Appl1
33	101.5	3.9	604	4	US-09-949-016-6031	Sequence 6031, Ap
34	101.5	3.9	604	5	US-09-689-366-4	Sequence 4, Appl1
35	101.5	3.9	604	5	PCT-US96-12860-4	Sequence 4, Appl1
36	101.5	3.9	613	4	US-09-949-016-10878	Sequence 10878, A
37	101.5	3.9	10182	3	US-09-134-001C-3159	Sequence 3159, Ap
38	101	3.8	1073	1	US-07-695-564-1	Sequence 1, Appl1
39	101	3.8	1073	1	US-08-241-387-1	Sequence 1, Appl1
40	101	3.8	1091	1	US-07-695-564-3	Sequence 3, Appl1
41	101	3.8	1091	1	US-08-241-387-3	Sequence 3, Appl1
42	101	3.8	1145	4	US-09-710-279-1432	Sequence 1432, Ap
43	101	3.8	1151	3	US-09-134-001C-3242	Sequence 3242, Ap
44	101	3.8	1359	4	US-09-134-000C-6374	Sequence 6374, Ap
45	100	3.8	449	4	US-09-107-433-4132	Sequence 4132, Ap

ALIGNMENTS

RESULT 1	US-09-118-442-11
Sequence 11, Appl1	Application US/09118442B
Patent No. 619361	
GENERAL INFORMATION:	
APPLICANT: Martino-Catt, Susan J.	
APPLICANT: Wang, Hongyu	
APPLICANT: Beach, Larry R.	
APPLICANT: Wang, Xun	
APPLICANT: Bowen, Benjamin A.	
TITLE OF INVENTION: Genes Controlling Phytate Metabolism in	
FILE REFERENCE: 0706	
CURRENT APPLICATION NUMBER: US/09/118,442B	
EARLIER FILING DATE: 1998-07-17	
EARLIER APPLICATION NUMBER: 60/055,446	
EARLIER FILING DATE: 1997-08-11	
EARLIER APPLICATION NUMBER: 60/055,526	
EARLIER FILING DATE: 1997-08-08	
EARLIER APPLICATION NUMBER: 60/053,944	
EARLIER FILING DATE: 1997-07-28	
NUMBER OF SEQ ID NOS: 31	
SOFTWARE: FastSeq for Windows Version 3.0	
SEQ ID NO 11	
LENGTH: 510	
TYPE: PRT	
ORGANISM: Zea mays	
US-09-118-442-11	
Query Match	89.2%; Score 2346; DB 3; Length 510;
Best Local Similarity	87.8%; Pred. No. 2e-222;
Matches	448; Conservative 28; Mismatches 34; Indels 0; Gaps 0;
QY	1 MEIENFKVESPVKXTETETIIVHNRNGTQIVPKSVNTOFKNTNHP 60
DB	1 MEIENFKVESPVKXTETETIIVHNRNGTQIVPKSVNTOFKNTNHP 60
QY	1 KLGWLVGMGNGSTLTGAVANREDISWATKDIQOANFGSLTQASIRVSGFOGE 120
DB	1 KLGWLVGMGNGSTLTGAVANREDISWATKDIQOANFGSLTQASIRVSGFOGE 120
QY	121 IYAPFKSLIPWNPDIYFGGWDISNMNADAMARAKYFDIDLOKQLRPYMESWPLPGI 180
DB	121 IYAPFKSLIPWNPDIYFGGWDISNMNADAMARAKYFDIDLOKQLRPYMESWPLPGI 180
QY	181 YDPDFIAANOEBRANVIGTKROVOQIIDKIAFKKATYDKVVLMTANTERYSLUV 240
DB	181 YDPDFIAANOEBRANVIGTKROVOQIIDKIAFKKATYDKVVLMTANTERYSLUV 240
QY	241 VGLNVTMENLAAVDRNBAETISPTLVAIACMENVPIINSPOVTFPGLIDLAIANT 300
DB	241 VGLNVTMENLAAVDRNBAETISPTLVAIACMENVPIINSPOVTFPGLIDLAIANT 300

```
Db 241 AGNDTMENTNLASVDKNEAEVSPSTLYAIACWMEGVPIGNSPQPTFVPGILDIAIKNNC 300
Qy 301 LIGGDDFKSGQTKKMSVLDVFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKKMSVLDVFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Qy 361 NVDDMVSNSAILYEPGHPDHVVVVKVVPVGDSCRAMDEYTSIEIFMGKNTVLHNTC 420
Db 361 NVDDMVSNSAILYEPGHPDHVVVVKVVPVGDSCRAMDEYTSIEIFMGKNTVLHNTC 420
Qy 421 EDSLIAAPIIDLVLLAEISTRIOFKAENEGKFSFHPVATITLSYTKAPLVPPTGPVNV 480
Db 421 EDSLIAAPIIDLVLLAEISTRIOFKAENEGKFSFHPVATITLSYTKAPLVPPTGPVNV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
```

RESULT 2

```
US-09-677-064-11
; Sequence 11, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Calc, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phycate Metabolism in
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-677-064-11
```

```
Query Match 89.2%; Score 2346; DB 3; Length 510;
Best Local Similarity 87.8%; Pred. No. 2e-222;
Matches 448; Conservative 28; Mismatches 34; Indels 0; Gaps 0;
```

```
Qy 1 MFIEFKVESPNVYKTEETETIOSVYNYETTELVHENRNGTYQVIKPKSVNTQFKTNTHVP 60
Db 1 MFIESFRVESHVRYGPEIESEYRYDTTELVEHKGDSARVVVPKSVKTNFRITRAVP 60
Qy 61 KLGWVLVGMGNGNSTLTGVIANREDISWATKDKIOQANYFGSLTQASAIRVSGFOGEE 120
Db 61 KLGWVLVGMGNGNSTLTGVIANREDISWATKDKIOQANYFGSLTQASAIRVSGSYNGEE 120
Qy 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKYFDIDLQKQLRPYMESWVPLPGI 180
Db 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADSMTRAKYLDIDLQKQLRPYMESWVPLPGI 180
Qy 181 YDPPFIANQSGRRANNVIKGTKEQVOQIIKDIKAFKATKVDKVVVLTMTANTERYSNLV 240
Db 181 YDPPFIANQSGRRANSVIKGTKEQVEQIIKDIRFKEKNKVDKIVLMTANTERYSNVC 240
Qy 241 VGLNDTMENTLAAVDNRNEAETSPSTLYAIACWMEVNPFIINGSPOPTFVPGILDIAIKNT 300
Db 241 VGLNDTMENTLAAVDKNEAEVSPSTLYAIACWMEGVPIGNSPQPTFVPGILDIAIKNNC 300
```

```
Qy 301 LIGGDDFKSGQTKKMSVLDVFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKKMSVLDVFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Qy 361 NVDDMVSNSAILYEPGHPDHVVVVKVVPVGDSCRAMDEYTSIEIFMGKNTVLHNTC 420
Db 361 NVDDMVSNSAILYEPGHPDHVVVVKVVPVGDSCRAMDEYTSIEIFMGKNTVLHNTC 420
Qy 421 EDSLIAAPIIDLVLLAEISTRIOFKAENEGKFSFHPVATITLSYTKAPLVPPTGPVNV 480
Db 421 EDSLIAAPIIDLVLLAEISTRIOFKAENEGKFSFHPVATITLSYTKAPLVPPTGPVNV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
```

RESULT 3

```
US-09-727-628-2
; Sequence 2, Application US/09727628
; Patent No. 6791013
; GENERAL INFORMATION:
; APPLICANT: Armstrong, Katherine
; APPLICANT: Hey, Timothy D
; APPLICANT: Folkerts, Otto
; APPLICANT: Smith, Kelley A
; APPLICANT: Hopkine, Nicole L
; TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
; FILE REFERENCE: 50597
; CURRENT APPLICATION NUMBER: US/09/727,628
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/168,612
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-727-628-2
```

```
Query Match 88.8%; Score 2337; DB 4; Length 510;
Best Local Similarity 87.8%; Pred. No. 1.5e-221;
Matches 448; Conservative 28; Mismatches 34; Indels 0; Gaps 0;
```

```
Qy 1 MFIEFKVESPNVYKTEETETIOSVYNYETTELVHENRNGTYQVIKPKSVNTQFKTNTHVP 60
Db 1 MFIESFRVESHVRYGPEIESEYRYDTTELVEHKGDSARVVVPKSVKTNFRITRAVP 60
Qy 61 KLGWVLVGMGNGNSTLTGVIANREDISWATKDKIOQANYFGSLTQASAIRVSGFOGEE 120
Db 61 KLGWVLVGMGNGNSTLTGVIANREDISWATKDKIOQANYFGSLTQASIRVSGSYNGEE 120
Qy 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKYFDIDLQKQLRPYMESWVPLPGI 180
Db 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMTRAKYLDIDLQKQLRPYMESWVPLPGI 180
Qy 181 YDPPFIANQSGRRANNVIKGTKEQVOQIIKDIKAFKATKVDKVVVLTMTANTERYSNLV 240
Db 181 YDPPFIANQSGRRANNVIKGTKEQVEQIIKDIRFKEKNKVDKVVVLTMTANTERYSNVC 240
Qy 241 VGLNDTMENTLAAVDNRNEAETSPSTLYAIACWMEVNPFIINGSPOPTFVPGILDIAIKNT 300
Db 241 AGNDTMENTNLASVDKNEAEISPTLYAIACVMEGVPIGNSPQPTFVPGILDIAIKNNC 300
Qy 301 LIGGDDFKSGQTKKMSVLDVFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKKMSVLDVFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSKEISKS 360
Qy 361 NVDDMVSNSAILYEPGHPDHVVVVKVVPVGDSCRAMDEYTSIEIFMGKNTVLHNTC 420
Db 361 NVDDMVSNSAILYEPGHPDHVVVVKVVPVGDSCRAMDEYTSIEIFMGKNTVLHNTC 420
```

421 EDSLLAAILDLVLLAELSTRIOFKAENEGKFSHPHVAATILSYLTKAPLVPGTVPVN 480
421 EDSLLAAILDLVLLAELSTRIOFKAENEGKFSHPHVAATILSYLTKAPLVPGTVPVN 480
481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
481 ALAKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 4

US-09-734-237B-73
; Sequence 73, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-73

Query Match 52.8%; Score 1389.5; DB 4; Length 534;
Best Local Similarity 53.1%; Pred. No. 4.9e-128;
Matches 278; Conservative 87; Mismatches 138; Indels 21; Gaps 8;

3 IENFVESPNNKYETEISQVYVNETTELHNNRGTQWIVKPKSVNYQFKNTHVP-K 61
9 ITSVAVVDKCTYKDNELLTKSYENA-VTKTASGRD--VTPVQDYVFKDLKREK 65
62 LGVNLVGGNGNSTLGGVIANREDISWATKDKIQANFYGSLTQASAIRVG-SFOGEE 120
66 LGIMLIGGNGNSTLVASVLANKNVFEQTEGVKQPNYFGSMTCSTLKGIDAEGND 125
121 IYAPKSLPMPNPPDIYFGGWDISNNMLADMAAPKFDIDLQKQLRPYMSMPLFGI 180
126 VYAPNSLLPMPVSPDFVSGWDINNADLYEAMQHSQVLEVDLQQLAKKSLVPLPSI 185
181 YDPDFIAANOBERANNVI-----KGT-----KQEQVOQIIKDIKAPKATKYDKVVLWMTA 231
186 YPDPFIAANOBERANNVCINLDEKGVNTRGKMTLQRLRRDIQNFKENALDKVILWMTA 245
232 NTERYSNLVGLNDTMENTMLAAVDNNEAISPSTLYAIACWENVPFINGSPONTFVGL 291
246 NTERVVEVSPGVNDTMENTMLQSIKNDHEIAPSTIFAASILEGVPIYNGSPQNTFVGL 305
292 IDLAIAKNTLLGGDDPKSGGQTRKMSVLDVFLVAGAGIKPTSVSYNHLGNNDGMULSAPO 351
306 VOLAEHEGTFIAGDDLKSGQTKLSVLAQFLVDAGIKPVSIAISYNHLGNNDGMULSAPO 365
352 FRSKSISKSNVVDVWVNSNALIYEP--GEHPDVVVIKYVVPVYVGSKRAMDEYSEIFMG 409
366 FRSKSISKSNVVDVWVNSNALIYEP--GEHPDVVVIKYVVPVYVGSKRAMDEYSEIFMG 425
410 GKNITVLNHTCEDSLAAILDLVLLAELSTRIOFKAENEGKFSHPHVAATILSY 464
426 GHNRSISHNVCEDSLAAILDLVLLAELSTRIOFKAENEGKFSHPHVAATILSY 485
465 YLTKAPLVPGTVPVNALSKORAMLENIMRACVGLAPENNMILE 508
466 YLTKAPLVPGTVPVNALSKORAMLENIMRACVGLAPENNMILE 508

RESULT 5

US-09-734-237B-75
; Sequence 75, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic protein derived from Saccharomyces cerevisiae myo-inosi
; OTHER INFORMATION: col-1-phosphate synthase, having a glycine residue inserted after
; OTHER INFORMATION: the initiating methionine
US-09-734-237B-75

Query Match 52.8%; Score 1389.5; DB 4; Length 534;
Best Local Similarity 53.1%; Pred. No. 4.9e-128;
Matches 278; Conservative 87; Mismatches 138; Indels 21; Gaps 8;

3 IENFVESPNNKYETEISQVYVNETTELHNNRGTQWIVKPKSVNYQFKNTHVP-K 61
10 ITSVAVVDKCTYKDNELLTKSYENA-VTKTASGRD--VTPVQDYVFKDLKREK 66
62 LGVNLVGGNGNSTLGGVIANREDISWATKDKIQANFYGSLTQASAIRVG-SFOGEE 120
67 LGIMLIGGNGNSTLVASVLANKNVFEQTEGVKQPNYFGSMTCSTLKGIDAEGND 126
121 IYAPKSLPMPNPPDIYFGGWDISNNMLADMAAPKFDIDLQKQLRPYMSMPLFGI 180
127 VYAPNSLLPMPVSPDFVSGWDINNADLYEAMQHSQVLEVDLQQLAKKSLVPLPSI 186
181 YDPDFIAANOBERANNVI-----KGT-----KQEQVOQIIKDIKAPKATKYDKVVLWMTA 231
187 YPDPFIAANOBERANNVCINLDEKGVNTRGKMTLQRLRRDIQNFKENALDKVILWMTA 246
232 NTERYSNLVGLNDTMENTMLAAVDNNEAISPSTLYAIACWENVPFINGSPONTFVGL 291
247 NTERVVEVSPGVNDTMENTMLQSIKNDHEIAPSTIFAASILEGVPIYNGSPQNTFVGL 306
292 IDLAIAKNTLLGGDDPKSGGQTRKMSVLDVFLVAGAGIKPTSVSYNHLGNNDGMULSAPO 351
307 VOLAEHEGTFIAGDDLKSGQTKLSVLAQFLVDAGIKPVSIAISYNHLGNNDGMULSAPO 366
352 FRSKSISKSNVVDVWVNSNALIYEP--GEHPDVVVIKYVVPVYVGSKRAMDEYSEIFMG 409
367 FRSKSISKSNVVDVWVNSNALIYEP--GEHPDVVVIKYVVPVYVGSKRAMDEYSEIFMG 426
410 GKNITVLNHTCEDSLAAILDLVLLAELSTRIOFKAENEGKFSHPHVAATILSY 464
427 GHNRSISHNVCEDSLAAILDLVLLAELSTRIOFKAENEGKFSHPHVAATILSY 486
465 YLTKAPLVPGTVPVNALSKORAMLENIMRACVGLAPENNMILE 508
467 YLTKAPLVPGTVPVNALSKORAMLENIMRACVGLAPENNMILE 508

RESULT 6

US-09-248-796A-17234
; Sequence 17234, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

```

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17234
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17234

Query Match          52.6%; Score 1384.5; DB 4; Length 525;
Best Local Similarity 52.0%; Pred. No. 1.5e-127;
Matches 266; Conservative 99; Mismatches 132; Indels 15; Gaps 5

QY      9 ESPNVKYTEIETIQSVYNYETTELVEHNNNGTYQWIVKPXSUNYOFTKHVTKGLVMVG 68
DB      13 KSSNSVTDKDDHLHTKYFYENS--VVEKDANG--KFIVTPTASPEEKVDLAKPVGILLVG 69
QY      69 WGNNGNSTLTGGVINREDISWATDCKIQANYPFSLTOQAIRG--SFOEEETIAPFK 126
DB      70 IGGNGTTLLGTATLDKNHISFEENEGGVKKPYYSVQTASTVKIGVKEGEDVYVPFN 129
QY      127 SLTPWNPPDDIFYGGMDSINNMILADAMARAKFDIDLQOLRPYMESWVLPGIYDPDFI 186
DB      130 SIVEPWNNDLVWDMDISGLPLDQAMRAKVLDTYLQOLPYLENKKPLESIYYPDFI 189
QY      187 AANOERANNVIK-----GTKOEYOQOIIKDIKAPEAKTVKRVVLMVMTANTERYSN 238
DB      190 ALNOSEKANNVFNQNGEYKTDNKMADEVKIRKDIRDEFAKNEIDKVIILMTANTERYAD 249
QY      239 LVVGLNDIMENILAAVDNREARISPTLYALACWENVPFINGSPONTFVPGTLIDLATAR 298
DB      250 VLPVNNDPADNLIKISKISHERIASTVPAVASILEKPYLINGSPONTFPVPGVIELAEKY 309
QY      299 NTLIGDDPKSGQOTQMKSVLVDFLYAGAIKPTSIVSYNHIGNDDMNLSAPOTFRSKSIS 358
DB      310 DSFIGDDPKSGQOTKIKSVLAQFLVDAGIKRPISTASYNHIGNDDMNSSPQOFPSKETS 369
QY      359 KSNVVDDMVNSNAIYY--EPGEHPHVVIKXPYPVGDSCRAMDEYTSEIFMGXNTITVL 416
DB      370 KQSVVDDIIESNELLYKNKSGBDKVDHCITYIKYLPVAVGSKVAMDYYSELMLGHNKISI 429
QY      417 HNTCEDSLAAPTIIDVLLABLSLRIOFKAENEGRHSFHVAATILSYLTAPLVPGQT 476
DB      430 HNVCEDSLAAPLIIDLVVATEFARVQVKGKSDYDELVELYVASILSYWLAKPLARPGEF 489
QY      477 PVNALSKORAMLNEMINRACVGLAPENNMIIE 508
DB      490 KPINGLNKOQLVNLISVYLGELPIDNELRFE 521

RESULT 7
US-09-902-540-12518
; Sequence 12518, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Goldan, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12518
```

```

; LENGTH: 444;
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12518

Query Match      20.1%; Score 528; DB 4; Length 444;
Best Local Similarity 29.4%; Pred. No. 3.5e-43;
Matches 135; Conservative 91; Mismatches 187; Indels 46; Gaps 11;

CY KLGWLVGWMGGNNGSTLTGGVIANREDISWATKDKIQQANYPGSLTQASAIRVGS-FQGE 119
DB KLAVLIPGLGANS-TTIMAGVELARKG-----KHPIGSLTGWGTARLGRKDR 61
CY EYAPFESKLLEPMVNEDDVFEGGMDTSNNQLADAMAPAKVPFDLDQOLRPYMESMPLPG 179
DB TV-KLNEIVPLAEKLDVAFGAMDIREDAEVAVRSGVLSDKHLEBVEKPFLOSIRKQG 119
CY IYDPPIANGERRANNVKGTR--QEQQOIIKOIKAREATKVVDKVVLMNTANTERYS 237
DB VHDPEFVRRIE---NHKATKTTHRESIEALRQDIRDKKEINATRAVNVVCSSVETFR 175
CY NLVVGELNPTMEMLLAADVRENAAIESPSTLYAICAWENPPINGSPO-NTFVPLDLAI 236
DB PLPSGFK-TLAFAEKALDENSPDINPTALYTAAIIEGPFNATATNASVDTPALQEMAK 234
CY ANTLIIGDDFEKSGGTOKMSVLVDPLVGAGIKPTSIVSYNHLGNNDGMNLISAPQTRSKE 356
DB QESVAVAGRDLKSQTMMKTVIAPALKAKMLGLDGWFSTNILGNRGDEVLDPAAFKAKE 294
CY ISKSVVUDMVNSNAIIVEGSEHPD----HVYVKVYVYDSKRAMEDEYSEIFMG 410
DB VTKSSVLDLTIL-----QPDLPELDYKKYAHKVALHYPRPDAGEGWNIDITGMVGY 347
CY KTTIVLHNTCEDSLAAPRIILDVLVLASTRIQFAENGKFSHPVATILSYLTRAP 470
DB PMQIKVNFCLCRSILAAPRLVIDIALFLDLAKLLEMKGIDEW-----MSFYFKTP 396
CY LVPGTPTPVNALISKOPAMLENIMTRAACVGIAPENNMLEY 509
DB MAQPGI.PVEHDLFIQTLTKMTLRVAGBDPTTHGLDY 435

RESULT 8
US-09-248-796A-20645
; Sequence 20645, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20645
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20645

Query Match      4.5%; Score 119.5; DB 4; Length 452;
Best Local Similarity 21.0%; Pred. No. 0.0071;
Matches 107; Conservative 80; Mismatches 158; Indels 165; Gaps 31;

CY PNVTXTET---EIGSVNYVE-TTELVA---ENRNNGTYQMIVPKSKVNYQPK----- 54
DB PSIVYPDSDITKOYQATNNSENATEETIYPILESK-----IVNIDANNYLKIITLOGVIAN 55
CY ---INTHTVPKLVN-LVWGNGNNGSTLTGGVIANREDISWATKDKIQQANY-FGSLTQA 108

```

```

Dh 56 HPTITINGQIMPLLIRPSSLTWSROSIEYIKRYIENLEITAFNIIDLSLATTPRGVQUTNS 115
Qy 109 SAIRVSGFOGEEIYAPFKSLPMPVNPDDIVFGMDISN-----MNL-----A 150
Dh 116 TVVVY---DDENI-----QIVPVVGYQALFKFAGKLKKNESGRTISREBKQNLPNLTSQOL 167
Qy 151 DAMARAKFPD-IDLOKQURPFWESNVPPLGIDYDPF-----IAAQOERANNVIGTKQ 203
Dh 168 EDLKNSDIFEVAVIIDOOGMWLDYIKDITKTNNEDNEFDVAKIVTENONGIPEAIISNTPT 227
Qy 204 BOVOQIINDI-KAFPEATK---VDKVV---VLMPTANERYSNLVGMDTMENTENLLAVDRN 257
Dh 228 EOQOQEOQDSNPKNELEKNYPTIDSKTOEKIWKG-ERFS---GTN-----N 270
Qy 258 EAEISPTLYALACWENYVPIINGSPOMTFVPLIDLAIRNTLIGADDPFKSGOTKWSV 317
Dh 271 LVKLISSTIYS--SLISPIDID-KRODCY-----DNIIIVGSIPTKTGLK-EAI 315
Qy 318 LV---DELYV-----GAGI-----KPTSIYSYHNLGNNDGNLISAPOTFRS 354
Dh 316 LILKNQDYLVEKPNALIDGSIINPQVNTALIKKYQOSSTINDYNEGSGDNNN----- 366
Qy 355 KEISKSNVDDWMSNALLIYEPGEHPDHVVYIKYVPYVGDSCRAMDE----- 401
Dh 367 ---SNSN-----SNSNOV-----PNSIKLYVYPPYPEMWKKPKKGGSGMHDVYFLGQ 411
Qy 402 -YTSIEFMG-----GKNTIYLANTCEDSL 425
Dh 412 IYSKOIYSGSSHGGKELFVSGDMYEEBVL 441

```

RESULT 9
 US-09-077-098A-6
 Sequence 6, Application US/09077098A
 Patent No. 6544519
 GENERAL INFORMATION:
 APPLICANT: TOKUNAGA, Ei-ji
 SAKAGUCHI, Masaaki
 MATSUO, Kazuo
 HAMADA, Fukuaburo
 TOKIYOSHI, Sachio
 TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS
 PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BROWDY AND NEWMARK
 STREET: 624 Ninth Street, N.W., Suite 300
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20001
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/077, 098A
 FILING DATE: 19-May-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP97/03222
 FILING DATE: 12-SEP-1997
 APPLICATION NUMBER: JP 27, 148/1996
 FILING DATE: 19-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: KORNEAU, Anne M.
 REGISTRATION NUMBER: 25, 618
 REFERENCE/DOCKET NUMBER: TOKUNAGA=1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-628-5197
 TELEFAX: 202-737-3528
 INFORMATION FOR SEQ ID NO: 6:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2042 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-077-098A-6

```

Query Match	4.5%	Score 118.5	DB 4	Length 2042
Best Local Similarity	23.0%	Pred. No. 0.11		
Matches 89	Conservative 53	Mismatches 146	Indels 99	Gaps 21

```

QY 37 NGTYOWIYVKRSVYQCKETHVHVKLGVMYKMGKNGNSSTLTGCVIANRED-----ISMA 91
Db 1160 SATOW--ANNGVV--KTNV-----LTTYSQNSTLIFG---MRBPSVKQIATG 1201
QY 92 TKDKIIOOANYEGSL-----TQASAIRGVSFOGEIYAPFESLLPMVVPDIDV---GWM 142
Db 1204 TYNTTGGANNQNLANTLTQOTTEATIGTSSVGSNTNAGFS-----LGADSVTPSKGAG 1255
QY 143 DISMNLTADAMARAKVPDIDLOKOLRPYMESMVPPLPGIYDPDPIAANOEBRANVI---- 198
Db 1259 TVKLSGVSDATADT---DAATLKQVKEKRTLLVG-----DMDITPAADSGGTSNGIFTYNL 1310
QY 199 ---KGTOEQVOOQIIOIKAFKAKETKVDKVVVLTANTANERSNLVGGNDTMENTLL-AAV 254
Db 1311 SLNKGTT-----VSATEKEVVSQKTYEALIRNATIGNIFTTGLDPTLTKINPNA 1355
QY 255 DRNEAESPTFLYIACVMENV-----PFIKSPQWTFVPGJLIDLIARNTLLIG 304
Db 1360 DQDLSNLSBESKNAITGLVDVVKKTNSPIYEPSTDSKKTIFVVG-VDFI---DTIREG 1411
QY 305 D--DFKSGQT-----KMSVLVDFLVAGIKPPTSIVSYNHLG---NDGMNLSAP 349
Db 1416 DATDCKLTTTSKSVESYVTNLANFSTDILLSDGRSNATLANGVGRRLRSDGFTIKSE 1475
QY 350 Q-TFRSKEIKSNVYDDMWSNALLYE 375
Db 1476 NFTLGSKOYNGS-----DSLGYMYD 1495

```

RESULT 10
 US-09-107-433-4707
 Sequence 4707, Application US/09107433
 Patent No. 6800744
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
 SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
 THERAPEUTICS
 NUMBER OF SEQUENCES: 5206
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: <Unknown>
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: <Unknown>
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,433
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/ 085131
 FILING DATE: May 12, 1998
 APPLICATION NUMBER: 60/051553
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:

```

; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-8277
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4707:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEITICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1..730
; SEQUENCE DESCRIPTION: SEQ ID NO: 4707:
US-09-107-433-4707
```

```

Query Match 4.4%; Score 116; DB 4; Length 730;
Best Local Similarity 20.8%; Pred. No. 0.035; Indels 165; Gaps 29;
Matches 108; Conservative 65; Mismatches 182;
```

```

QY 90 WATKDKIQANYFG-----SLTQSAIRVGSFGGEIYAFPKSLPMVNPDDIVFGMD 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 218 WGVEDASKK--YFGVSASEVSLDQAATL-AGMLKGPGLYNPLNSVEDSTNRDITV----- 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 144 ISNNMLA---DAMARAKVFDIDLQKQLRPYMESWV---LPGIVDPDFIAA-----NOE 191
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 270 LQNVVAAGYIDKNOETEAEDVMTSQLDKYEGKISDYRPSYFDVAVNEAVSKYNLTBE 329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 192 ERANN-----VIKTKQEQVOQIITKDIKAFKEA-----TKVDKVVYLM 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 330 EIVANGRIYIELDQNYQAMQIYENISLPPRAEDGTPAQSGSVALPEPTKGVRGVVQ 389
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 230 TANTER-----YSN-----LVV-----GLNDTMENLLAANDRNAEI 261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 390 VADNDKGTGFNFNVAATQSKSPGSTIKPLVYTPAVEAGMALNKQLDNHTMQYDSYKVD- 448
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 262 SPSTLYIACMENVPFINGSPONTFVPGLI---DLAIARNTLIGDDPFKSGQTKMSYL 318
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 449 ---NYAGIKTSREVPMTQALAESLNLPAVATVVDLGVDK-AFEAGEKFGMLMEKVDRLV 503
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 319 VDFLVGAGIK--PTSIIVSYNHLGNNDGM-----NISAPQTFRSKEISKSNNVDD 365
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 504 -GVALGSGVETNPLQMAQVAAFAFNEGIMPEAHFISRIENASG-QVIANKHNSQKRVYIDK 561
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 366 MV-----NSNAIIYEPGEHPDHVVVIKYVYVGDSSKRAMD-EYTSIEIFMG- 410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 562 SVADKMTSMMLGTFTNGTGISSSPA---DYVMAGK---TGTEAVNPETYSQWVIGY 614
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 411 -KNITVLH----NTCEDSLAAPILIIDLVLIELSTRIOFKANEKGFHSFHPVA-TIL 463
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 615 TPDVVISHWLGFPTTDENHYLAG-----STSGAAHYFRNIANTIL 655
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 464 SYLTAKPLVPPTGPVNVNLSKORAMLENIMRACVGLAPEN 503
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 656 -----PYTPGSTFTVENAYKQN-----GIAPAN 678
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```

RESULT 11
US-09-583-110-4720
; Sequence 4720, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
```

```

; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4720
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4720
```

```

Query Match 4.4%; Score 116; DB 4; Length 731;
Best Local Similarity 20.8%; Pred. No. 0.035; Indels 165; Gaps 29;
Matches 108; Conservative 65; Mismatches 182;
```

```

QY 90 WATKDKIQANYFG-----SLTQSAIRVGSFGGEIYAFPKSLPMVNPDDIVFGMD 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 215 WGVEDASKK--YFGVSASEVSLDQAATL-AGMLKGPGLYNPLNSVEDSTNRDITV----- 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 144 ISNNMLA---DAMARAKVFDIDLQKQLRPYMESWV---LPGIVDPDFIAA-----NOE 191
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 267 LQNVVAAGYIDKNOETEAEDVMTSQLDKYEGKISDYRPSYFDVAVNEAVSKYNLTBE 326
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 192 ERANN-----VIKTKQEQVOQIITKDIKAFKEA-----TKVDKVVYLM 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 327 EIVANGRIYIELDQNYQAMQIYENISLPPRAEDGTPAQSGSVALPEPTKGVRGVVQ 386
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 230 TANTER-----YSN-----LVV-----GLNDTMENLLAANDRNAEI 261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 387 VADNDKGTGFNFNVAATQSKSPGSTIKPLVYTPAVEAGMALNKQLDNHTMQYDSYKVD- 445
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 262 SPSTLYIACMENVPFINGSPONTFVPGLI---DLAIARNTLIGDDPFKSGQTKMSYL 318
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 446 ---NYAGIKTSREVPMTQALAESLNLPAVATVVDLGVDK-AFEAGEKFGMLMEKVDRLV 500
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 319 VDFLVGAGIK--PTSIIVSYNHLGNNDGM-----NISAPQTFRSKEISKSNNVDD 365
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 501 -GVALGSGVETNPLQMAQVAAFAFNEGIMPEAHFISRIENASG-QVIANKHNSQKRVYIDK 558
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 366 MV-----NSNAIIYEPGEHPDHVVVIKYVYVGDSSKRAMD-EYTSIEIFMG- 410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 559 SVADKMTSMMLGTFTNGTGISSSPA---DYVMAGK---TGTEAVNPETYSQWVIGY 611
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 411 -KNITVLH----NTCEDSLAAPILIIDLVLIELSTRIOFKANEKGFHSFHPVA-TIL 463
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 612 TPDVVISHWLGFPTTDENHYLAG-----STSGAAHYFRNIANTIL 652
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 464 SYLTAKPLVPPTGPVNVNLSKORAMLENIMRACVGLAPEN 503
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 653 -----PYTPGSTFTVENAYKQN-----GIAPAN 675
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```

RESULT 12
US-09-489-039A-11570
; Sequence 11570, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11570
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11570
```


Query Match 4.4%; Score 115; DB 4; Length 842;
Best Local Similarity 24.7%; Pred. No. 0.05;
Matches 97; Conservative 51; Mismatches 134; Indels 110; Gaps 25;

QY 93 KDKIOANVFGSLTQASAIRVGSFQGEIYAPFKSLLPVNPDPDVFQGMDSNM--NLA 150
DB 460 REGTQRFDFENLRVAGQY-IGSFR---TMTPTASIQQYVS-----DLKQVGNCE 505
QY 151 DAMARAKYVD-----IDLQKQLRPYMESWVPLPGIY-----DPDFIANQGEER--A 194
DB 506 GLAGRAAAAFDHVQAQAQRISQQRSDPF-QSAVDI-GAYKPISNSTPPDAIAEYVNRVAA 563
QY 195 NNVIKQ-----TKQEQVOQIIKDIKAFKATKVDKVVVMTANTERYSLVVGANDT 246
DB 564 QDLVAIGITPEPLSKQE--SQVLTD--AVRNSTVNCQAI-----SLQGLGRT 608
QY 247 M-ENLLAIVDRNEAISPSTLYAIACV-MENVPFINGS---PONTFV---PGLIDLAITAR 298
DB 609 LPPKALRVSASIASGSGTAYALLLGQDQNDNRSGIIRYSQFVSYKPTLDKRYDAK 668
QY 299 NTLIGGDDFKSGQTKKSVLVDFLVAGAIKPTISVSYNHL-----GNNDGMNLSAPQ 350
DB 669 -TVLAGDQMLNPTKAMKD-----AGISAVSIPSEDEKTKYFDKKEGNARFVSAQAARQ 719
QY 351 -----TRSKKISKSNNV--DDMVNSAALLYEPCG-----HPDHVVVIKTVP 390
DB 720 VAMGNFRSAYALAAVQSDGASKTNTVSPSDIAEKRAQYATGVYKGLGSDVV-----MP 775
QY 391 YVGSKRAMDEYTS---EIFWGGKNTIYLAH 418
DB 776 FGMDKTTFRDRYTLAAGREAMTQAGLNPSSLEN 807

RESULT 13

US-09-014-897-2
Sequence 2, Application US/09014897
Patent No. H002085

GENERAL INFORMATION:

APPLICANT: Hoskins, Johann
Jaskunas, S. Richard
Rockey, Pamela K.
Zhao, Genshi
Rostock, Paul R. Jr.
No. H002085tis, Franklin H.
TITLE OF INVENTION: Penicillin Binding Protein From
Streptococcus Pneumoniae
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,897
FILING DATE: 28-Jan-1998
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/731,716
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-10,887
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 731 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-014-897-2

Query Match 4.3%; Score 114; DB 1; Length 731;
Best Local Similarity 20.8%; Pred. No. 0.05;
Matches 108; Conservative 64; Mismatches 183; Indels 165; Gaps 29;

QY 90 WATDKIOQANVFG-----SLTQASAIRVGSFQGEIYAPFKSLLPVNPDPDVFQGM 143
DB 215 WGVEDASKK--YFGVSASEVSLDQAAVL--AGMLKGPGLYNPLNVSDESTNRDDTV----- 266
QY 144 ISNNMLA---DAMARAKYFDIDLQKQLRPYMESWV---LPQIYDPDFIAA-----NOE 191
DB 267 LQNVYAAGYIDKQGETEALEVDMTSQLDKTEGKISDRYRSYFPAVVNNAVSKYNLTBE 326
QY 192 ERANN-----VIKGTQEQVOQIIKDIKAFKA-----TKVDKVVYLA 229
DB 327 EIVNNGYRIYTELDQNVQANQIYVENTSLFPRAEDGTFAGSGVALBPKTGVAGVVGQ 386
QY 230 TANTER-----YSN-----LVY-----GLNDMENLLAIVDRNEAEI 261
DB 387 VADNDKTFRRNFNATQSRGSGSTIKPLVYTTPAVEAGMALNKQLDHRTMQDYSKYVD- 445
QY 262 SPSTLYAIACWENVPFINGSFONTFVPLGL---DLAIAARNTLIGGDDFKSGQTKKSVL 318
DB 446 ---NYAGIKTSREVPMTQSLAESLNLPAVATVNDLVGDK-AFEAGEKFGGLMEKVDKYL 500
QY 319 VDFLVAGIK--PTISVSYNHLGNNDGM-----NLSAPQTFRSKEISKSNNVDD 365
DB 501 -GVMLGSGVETNPLOMAQAYAAFAFNBGLMPEAHFISRIENASG-QVIAISHKNSQKRVLDK 558
QY 366 MV-----NSNAIYYEGEHPDHVYVIVYVGSKRAMD-ETISELPMKG- 410
DB 559 SVADKMTSMMLGTFTNGTGISSSPA---DYWAGK---TGTEVAFNPEYTSDDQVIGY 611
QY 411 -KNRTIVLH-----NTCEPSLLAAPRIIDLVLLAELSTRIOFKAENEGKFSHPVA-TIL 463
DB 612 TPDVVISHWLGPPTDENHYLAG-----STNSGAHVFRNIANTITL 652
QY 464 SYLTCAPLVPPGTVPVNAISKQRALEINMRACVGLADEN 503
DB 653 -----PYTPGSTFTVENAYAKQN-----GIADAN 675

RESULT 14

US-08-731-716-2
Sequence 2, Application US/08731716
Patent No. 5789202

GENERAL INFORMATION:

APPLICANT: Hoskins, Johann
Jaskunas, S. Richard
Rockey, Pamela K.
Zhao, Genshi
Rostock, Paul R. Jr.
No. 5789202tis, Franklin H.
TITLE OF INVENTION: Penicillin Binding Protein From
Streptococcus Pneumoniae
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/731,716
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-10,887
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 731 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-731-716-2

Query Match 4.3%; Score 114; DB 1; Length 731;
Best Local Similarity 20.8%; Pred. No. 0.055;
Matches 108; Conservative 64; Mismatches 183; Indels 165; Gaps 29;

QY 90 WATDKIQOANYFG-----SITQASAIRVGSFQGEIYAPFKSLIPMNPDPIYFGWD 143
DB 215 WGVEDASKK-YFGVASEVSLDQAAATL-AGMLKGPDLNPLNSVEDSTNRDVT----- 266
QY 144 ISNNMLA---DAMNAKVFDDIDLOKOLRPYMESWP---LPGIYPPDIAA-----NOE 191
DB 267 LQNNAAAYIDKNQDTEAEVDMTSOLHDKYEGKISDYRPSYFPAVVNEAVSKYNLTKE 326
QY 192 ERANN-----VIKGTKEOVQOIIKDIKAFKEA-----TKVDKVVILW 229
DB 327 EIVNNGVRIYELDQNYQANNQIYVENTSLFPRADGTFAGSGVALPKTGVAGVVGQ 386
QY 230 TANTER-----YSN-----LVV-----GLNDTMENLLAVDENEAEI 261
DB 387 VADNDKTFRPNFYATQGRSPGSTIKPLVYTPAVEAGMALNKQLDNHTMQYDSYKVD 445
QY 262 SPSTLYAIACWENVPFINGSPQNTFVPGLI---DLAARTLLIGDGFQSKGQTFMKSVL 318
DB 446 ---NYAGIKTSREVPMTQSLAESLNLPAVATVNDLGVDK-AFEXGKFGNMEKVDKVL 500
QY 319 VDFIVGAGIK--PTSIVSYNHLGNNDGM-----NLAPQTFRSKEISKSNNVVD 365
DB 501 -GVALGSGVEINPLQMAQAYAAFNIEGLMPEAHFISRIENASG-QVISHKNSQKRVIDK 558
QY 366 MV-----NSNAIIYEPGEHPDHVVVITVPYVGDSSKRAMD-EYTSIEFMGG- 410
DB 559 SVADKMTSMMLGTFTNGTGISSSPA---DYWAGK---TGTTAEVFNPEYTSIQWVIGY 611
QY 411 -KNTIVLH-----NICEDESLAARPIILDLVLLAELSTRIOFKAENEGKFSFHPVA--TIL 463
DB 612 TPDVVISHMLGPTDENHYLAG-----STNGAAHVFRNLIANTIL 652
QY 464 SYLTAPLVPPGTPYVNALSKORAMLENIMRACVGIADEN 503
DB 653 -----PYTPGSTPTVENAYKON-----GIADAN 675

RESULT 15
US-09-489-039A-8455
Sequence 8455, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 8455
LENGTH: 778
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8455

Query Match 4.3%; Score 113.5; DB 4; Length 778;
Best Local Similarity 19.3%; Pred. No. 0.068;
Matches 78; Conservative 61; Mismatches 149; Indels 117; Gaps 16;

QY 68 GWGNGNSTLTGVIANREDISWATKDIQO-----ANYFSLTQASAIRV----- 113
DB 160 GMRGDRDS-----RGDTSWVPAEMIDHIDVIRGPAAARYNGAMGVNIVTKPTT 210
QY 114 ---GSF-----QGEIYAPFKSLPMPNP--DDIVFGWMDISNNMLADAM----- 153
DB 211 PEWHGSMWYNNAPDQHRREGATKRTNFSLNGPLSDSVSFNLMGNLSKTQADADINAGHE 270
QY 154 -----ARAKVFDIDLQOLRPYMESWVPLPGIYDPDFIANQGEERANNVYKG 200
DB 271 AERTGSYAGSYPAAREGVNNDIHSKLRMEFAPQAL-----EFEAG--YSRQGNLYAG 322
QY 201 -TKQEQVQOIIKDIKAFKEATKVDKV--VLMTANTE-----RYSNLVVGJ 243
DB 323 DTQNTNSTVLVKSMTYG--KETRRLYRQTYGVWTGMDNGVTSNSYAQYEHTRNSRMDEGL 381
QY 244 NDTEMENLLAADNRBAEISPELTLYAIACWENVPFINGSPQNTFVPGLIIDLARTLLIG 303
DB 382 AGTEGIFSSSEFSDIDLADVLHSEV---NIPFTLGVDON-----LTL 422
QY 304 GDPEKSGQTKKKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNLAPQTFRSKEISKSNNV 363
DB 423 CTEN--NQGRMKD-----GVSTQALSY---GIIDGSAIRGRSPYSABEIFSLFTE 468
QY 364 DMVNSNAIIYEPGEHPDH--VVVITVPYVGDSSKRAMDEYTSI 406
DB 469 DNMALTDSTMLTLPALRPDHSIIVGNMNSPILSLQELTDWTTLKL 513

Search completed: June 7, 2005, 16:48:55
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 15:09:19 ; Search time 92.5 Seconds

(without alignments)
2132.409 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632

Sequence: 1 MREHPKVECPNVKYTEI.....NIMRACVGLAPENNILEYK 510

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2632	100.0	510	2	AAW9740 Soybean w
2	2632	100.0	510	8	ADQ14499 Mutant so
3	2632	100.0	510	8	ADQ14491 Wild type
4	2632	100.0	510	8	AD882000 Soybean m
5	2632	100.0	510	8	AD881994 Soybean m
6	2637	99.8	510	2	AAW9741 Soybean m
7	2637	99.8	510	8	ADQ14495 Mutant so
8	2637	99.8	510	8	AD881998 Soybean m
9	2600	98.8	510	8	ADQ14503 Mutant so
10	2600	98.8	510	8	ADQ14505 Wild type
11	2600	98.8	510	8	AD882004 Soybean m
12	2600	98.8	510	8	AD882006 Soybean m
13	2593	98.5	510	8	ADQ14501 Mutant so
14	2593	98.5	510	8	AD882002 Soybean m
15	2472	93.5	536	2	AA124477 Nicotiana
16	2432	92.4	505	8	ADG73738 Myo-inos
17	2378	90.3	510	8	AA650575 Arabidops
18	2378	90.3	510	8	ADN73525 Thale cre
19	2378	90.3	581	3	AA650574 Arabidops
20	2378	90.3	645	3	AA650573 Arabidops
21	2368	90.0	510	4	AA688935 Brassica
22	2352.5	89.4	511	3	AA690861 Arabidops
23	2352.5	89.4	511	3	AA632501 Arabidops
24	2352.5	89.4	534	3	AA690860 Arabidops
25	2344	89.1	510	2	AAW96259 Phytate p

26	2340	88.9	510	8	ADP43920 Rice L-my
27	2336	88.8	510	2	AAW9740 Maize myo
28	2335	88.7	510	4	AA647286 MIP synth
29	2080	79.0	446	3	AA632502 Arabidops
30	2080	79.0	446	3	AA609862 Arabidops
31	1871	71.1	512	8	ADP43919 Porterea
32	1768	67.2	380	3	AA632503 Arabidops
33	1607	61.1	558	4	AA65577 Human hMT
34	1607	61.1	558	4	AA640076 Human pol
35	1607	61.1	558	4	AA692849 Human pro
36	1607	61.1	558	4	AA693732 Human pro
37	1607	61.1	558	4	AA618891 Human ova
38	1596.5	60.7	534	6	ABJ25857 Aspergill
39	1592.5	60.5	557	4	AA65576 Murine mH
40	1592.5	60.5	534	6	ABJ26457 Aspergill
41	1585	60.2	565	4	AB655497 Drosophill
42	1556.5	59.1	659	8	ADN99714 Novel hum
43	1525	57.9	505	8	ADG73737 Aspergill
44	1525	57.9	532	8	ADG73736 Aspergill
45	1420	54.0	504	7	ADB64735 Human pro

ALIGNMENTS

RESULT 1	
ID	AAW9740 Standard; protein; 510 AA.
XX	
AC	AAW9740;
XX	
DT	17-OCT-2003 (revised)
DT	02-FEB-1999 (first entry)
XX	
DE	Soybean wild-type myo-inositol 1-phosphate synthase.
XX	
KW	Soybean, myo-inositol 1-phosphate synthase; raffinose; stachyose;
KW	phytic acid.
XX	
OS	Glycine max; line LR13.
XX	
PN	MO9645448-A1.
XX	
PD	15-OCT-1998.
XX	
PF	07-APR-1998; 98WO-US006822.
XX	
PR	08-APR-1997; 97US-00835751.
XX	
PA	(DUPO) DU PONT DE NEMOURS & CO E I.
XX	
PI	Hitz WD; Sebastian SA;
XX	
DR	WPI: 1998-568353/48.
XX	
DR	N-PSDB; AAW62440.
XX	
PT	Soybean plants containing altered myo-inositol 1-phosphate gene - useful
PT	for generating plants with altered levels of e.g. raffinose, stachyose,
XX	phytic acid, etc.
XX	
PS	Example 5; Page 45-47; 63pp; English.
XX	
CC	This is the amino acid sequence of soybean myo-inositol 1-phosphate
CC	synthase (MI 1-PS) deduced from the coding region of an isolated cDNA
CC	clone (see AAW62440) . MI 1-PS is involved in glucose metabolism to phytic
CC	acid, raffinose and stachyose. A mutant MI 1-PS (see AAW9741) has been
CC	identified in soybean line LR33, a mutagenised line of low raffinose
CC	saccharide phenotype. Sequencing revealed a single base change mutation
CC	in the LR33 gene sequence that resulted in a K396N substitution in the
CC	mutant protein. The mutation results in a seed phenotype of very low
CC	raffinose saccharide sugars, very high sucrose and low phytic acid. The
CC	mutated nucleic acid is used to alter the raffinose saccharide, sucrose,
CC	phytic acid and inorganic phosphate content of soybean seeds, leading to

CC useful soybean products, e.g. a seed phytic acid content of less than 17
CC ug/g, a seed content of raffinose and stachyose combined of less than
CC 14.5 ug/g, and a seed sucrose content greater than 200 ug/g. (Updated on
CC 17-OCT-2003 to standardise OS field)

XX Sequence 510 AA;

Query Match 100.0%; Score 2632; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1,2e-212;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIENFKVECPNVKKTETETISQVYVETTELHVENRNGTYQWIVKPKSVKYEFTKNIHVP 60
DB 1 MEIENFKVECPNVKKTETETISQVYVETTELHVENRNGTYQWIVKPKSVKYEFTKNIHVP 60
QY 61 KLGWVLVGMGNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
DB 61 KLGWVLVGMGNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPWVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPWVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEBRANNVIKGTQOEVOQIIDIKAFKATKYDKVVLMTANTERYSNLV 240
DB 181 YDPDFIAANOEBRANNVIKGTQOEVOQIIDIKAFKATKYDKVVLMTANTERYSNLV 240
QY 241 VGLNDTMENLLAADVNRNEAISPSTLYAIACWENVPFINSPOPTFVGLIDLAIAANT 300
DB 241 VGLNDTMENLLAADVNRNEAISPSTLYAIACWENVPFINSPOPTFVGLIDLAIAANT 300
QY 301 LIGGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
QY 361 NVVDDMVNSNALIYEPGEHPHVVIKYVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC 420
DB 361 NVVDDMVNSNALIYEPGEHPHVVIKYVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC 420
QY 421 EESLLAAPTILDLVLAELSTRIEFKAENEGKFSHPATILSYTKAPLVPGTPPVN 480
DB 421 EESLLAAPTILDLVLAELSTRIEFKAENEGKFSHPATILSYTKAPLVPGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 2

ADQ14499
ID ADQ14499 standard; protein; 510 AA.

XX ADQ14499;

DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #2.

KW Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
KW mutant; mutain.

OS Glycine max.
OS Synthetic.

PN US2004128713-A1.

PD 01-JUL-2004.

PF 21-NOV-2003; 2003US-00718952.

PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.

PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX WPI: 2004-53135/51.
XX N-PSDB; ADQ14498.

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

PS Example 8; SEQ ID NO 10; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant,
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.

SO Sequence 510 AA;

Query Match 100.0%; Score 2632; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 1,2e-212;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIENFKVECPNVKKTETETISQVYVETTELHVENRNGTYQWIVKPKSVKYEFTKNIHVP 60
DB 1 MEIENFKVECPNVKKTETETISQVYVETTELHVENRNGTYQWIVKPKSVKYEFTKNIHVP 60
QY 61 KLGWVLVGMGNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
DB 61 KLGWVLVGMGNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPWVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPWVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEBRANNVIKGTQOEVOQIIDIKAFKATKYDKVVLMTANTERYSNLV 240
DB 181 YDPDFIAANOEBRANNVIKGTQOEVOQIIDIKAFKATKYDKVVLMTANTERYSNLV 240
QY 241 VGLNDTMENLLAADVNRNEAISPSTLYAIACWENVPFINSPOPTFVGLIDLAIAANT 300
DB 241 VGLNDTMENLLAADVNRNEAISPSTLYAIACWENVPFINSPOPTFVGLIDLAIAANT 300
QY 301 LIGGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
QY 361 NVVDDMVNSNALIYEPGEHPHVVIKYVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC 420
DB 361 NVVDDMVNSNALIYEPGEHPHVVIKYVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC 420

QY 421 EDSLAAPIILDLVLLAEISTRIEFAKNEGKFSHPVATILSYTKAPLVPGTPVNV 480
 DB 421 EDSLAAPIILDLVLLAEISTRIEFAKNEGKFSHPVATILSYTKAPLVPGTPVNV 480
 QY 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
 DB 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510

RESULT 3
 ADQ14491
 ID ADQ14491 standard; protein; 510 AA.
 XX ADQ14491;
 AC
 XX 23-SEP-2004 (first entry)
 DT
 XX Wild type soybean myo-inositol 1-phosphate synthase polypeptide #1.
 DE
 XX Soybean: myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
 KM raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.
 XX Glycine max.
 OS
 XX US2004128713-A1.
 PN
 XX 01-JUL-2004.
 PD
 XX 21-NOV-2003; 2003US-00718952.
 PF
 XX 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 PR 26-APR-1999; 99US-00299315.
 PR 11-MAR-2002; 2002US-00025003.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PA (GRAC/) GRACE D J.
 PA (STRE/) STREIT L G.
 PA
 PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
 PI
 DR WPI: 2004-533135/51.
 DR N-PSDB; ADQ14490.
 XX
 PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytic acid and increased sucrose, leading to valuable and useful soybean
 PT products.
 XX
 XX Claim 3; SEQ ID NO 2; 48bp; English.

Query Match 100.0%; Score 2632; DB 8; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1.2e-212;
 Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIENKVECPNVKKTETETIQSVVYVETTELVHNRNGTYQYIVKPKSVKYEFTNIHVP 60
 DB 1 MEIENKVECPNVKKTETETIQSVVYVETTELVHNRNGTYQYIVKPKSVKYEFTNIHVP 60
 QY 61 KLGVMILVGMGANGSTLTGCVIANKREGISWATKDKIQQANYFGSLTQASAIRVGSFGSE 120
 DB 61 KLGVMILVGMGANGSTLTGCVIANKREGISWATKDKIQQANYFGSLTQASAIRVGSFGSE 120
 QY 121 IYAPFKSLIPMNPDDIVFGGWDISNMNLADMAARKVFDIDLOKQLRPYMESMLPLPGI 180
 DB 121 IYAPFKSLIPMNPDDIVFGGWDISNMNLADMAARKVFDIDLOKQLRPYMESMLPLPGI 180
 QY 181 YPDPFIAANQERRANNVIKGTQBOVOQIIKDIIKAFKATKYDKVVLWTANTERTSYNLV 240
 DB 181 YPDPFIAANQERRANNVIKGTQBOVOQIIKDIIKAFKATKYDKVVLWTANTERTSYNLV 240
 QY 241 VGLNDTMENTLAAVDNREAEISPTLYAIAQWENVPTNGSPONTFVPLIDLAIRNT 300
 DB 241 VGLNDTMENTLAAVDNREAEISPTLYAIAQWENVPTNGSPONTFVPLIDLAIRNT 300
 QY 301 LIGDPDFKSGQTKMSVLVDFVAGIKPTSTIVSYNHIGNNQNMLSAPOTRSKESKS 360
 DB 301 LIGDPDFKSGQTKMSVLVDFVAGIKPTSTIVSYNHIGNNQNMLSAPOTRSKESKS 360
 QY 361 NVVDDMVNSNALIYFGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
 DB 361 NVVDDMVNSNALIYFGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
 QY 421 EDSLAAPIILDLVLLAEISTRIEFAKNEGKFSHPVATILSYTKAPLVPGTPVNV 480
 DB 421 EDSLAAPIILDLVLLAEISTRIEFAKNEGKFSHPVATILSYTKAPLVPGTPVNV 480
 QY 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
 DB 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510

RESULT 4
 ADS82000
 ID ADS82000 standard; protein; 510 AA.
 XX
 AC ADS82000;
 XX
 XX 18-NOV-2004 (first entry)
 DT
 XX Soybean myo-inositol 1-phosphate synthase wild-type 3.
 DE
 XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
 KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
 XX
 OS Glycine max; line 29004JP01.
 PN
 XX US2003074685-A1.
 PD
 XX 17-APR-2003.
 XX
 PF 11-MAR-2002; 2002US-00025003.
 PR
 XX 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PA
 PI Hitz WD, Sebastian SA;
 PI
 DR WPI: 2004-639957/62.
 DR N-PSDB; ADS81999.
 XX

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

PS Example 8; SEQ ID NO 10; 34pp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC heritable protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.

CC Sequence 510 AA;

Query Match 100.0%; Score 2632; DB 8; Length 510;

Best Local Similarity 100.0%; Pred. No. 1.2e-212; Mismatches 0; Gaps 0;

Matches 510; Conservative 0; Indels 0; Gaps 0;

QY 1 MFENFKVCEPNVKTETETIQSVNYYETTELHNNNGTYQWIVKPKSVKFEKTNINHP 60
DB 1 MFIENFKVCEPNVKTETETIQSVNYYETTELHNNNGTYQWIVKPKSVKFEKTNINHP 60
QY 61 KLGWLVGVGNNNGSTLTGCVIANREGISWATKDKIQQANYGSLTQASAIRVGSFQGE 120
DB 61 KLGWLVGVGNNNGSTLTGCVIANREGISWATKDKIQQANYGSLTQASAIRVGSFQGE 120
QY 121 IYAPFKSLIPMNPPDIVFGWMDISNMNLADAMAAKAVDIDLOKQLRPMESMLPLPGI 180
DB 121 IYAPFKSLIPMNPPDIVFGWMDISNMNLADAMAAKAVDIDLOKQLRPMESMLPLPGI 180
QY 181 YDPDFIAAQBERRANNVIGKTQEQVQOIIKDKIKAFKATKYDKVVLMTANTERTSNLY 240
DB 181 YDPDFIAAQBERRANNVIGKTQEQVQOIIKDKIKAFKATKYDKVVLMTANTERTSNLY 240
QY 241 VGLNDTMENLLAADVBNENAEISBSTLYAIACMENVPFINSBPONTFVGLIDLAIRNT 300
DB 241 VGLNDTMENLLAADVBNENAEISBSTLYAIACMENVPFINSBPONTFVGLIDLAIRNT 300
QY 301 LIGGDFKSGQTKMSVLDVFLVAGIKRSTSVSYNHLGNNGMNLAPQTRSRKISKS 360
DB 301 LIGGDFKSGQTKMSVLDVFLVAGIKRSTSVSYNHLGNNGMNLAPQTRSRKISKS 360
QY 361 NVVDDMVNSNALIYERGEHPDHVVVVKYVPYVGSKRADMETSEIFMGKSTIYLAHNC 420
DB 361 NVVDDMVNSNALIYERGEHPDHVVVVKYVPYVGSKRADMETSEIFMGKSTIYLAHNC 420
QY 421 EDSLAAPIILDVLIAELSTRIEFKAENEGKRFHSFHPVATILSYTKAPLVPCTPVVN 480
DB 421 EDSLAAPIILDVLIAELSTRIEFKAENEGKRFHSFHPVATILSYTKAPLVPCTPVVN 480

DB 421 EDSLAAPIILDVLIAELSTRIEFKAENEGKRFHSFHPVATILSYTKAPLVPCTPVVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510

RESULT 5
ADS81994
ID ADS81994 standard; protein; 510 AA.

ADS81994;

18-NOV-2004 (first entry)

Soybean myo-inositol 1-phosphate synthase wild-type 1.

KW Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid,
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.

OS Glycine max; cultivar Wye.

PN US2003074685-A1.

PD 17-APR-2003.

PF 11-MAR-2002; 2002US-00025003.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.

PA (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PI Hitz WD, Sebastian SA;

DR WPI; 2004-639957/62.

DR N-PSDB; ADS81993.

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

PS Claim 3; SEQ ID NO 2; 34pp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence

CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.
XX

Sequence 510 AA;

Query Match 100.0%; Score 2632; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.2e-212;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MEIENFKVCEPNVKTETETISQVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
DB 1 MEIENFKVCEPNVKTETETISQVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
QY 61 KLGWNLVGMGNGNSTLTGCVIANREGISMAFKDKIQOANYGSLTQASAIRVGSFGQGE 120
DB 61 KLGWNLVGMGNGNSTLTGCVIANREGISMAFKDKIQOANYGSLTQASAIRVGSFGQGE 120
QY 121 IYAPFKSLIPMVPNDIVFGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPMVPNDIVFGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDTIAANQEBRANNVIKGTQEQVOQIIKDIFKFKATKVDKVVVLTANTERYSNLV 240
DB 181 YDPDTIAANQEBRANNVIKGTQEQVOQIIKDIFKFKATKVDKVVVLTANTERYSNLV 240
QY 241 YDPDTIAANQEBRANNVIKGTQEQVOQIIKDIFKFKATKVDKVVVLTANTERYSNLV 240
DB 241 YDPDTIAANQEBRANNVIKGTQEQVOQIIKDIFKFKATKVDKVVVLTANTERYSNLV 240
QY 241 VGLNDTMENTLAAVDRNEABISPSLTLYAACWENVPFINSPOPTFVGLIDLAIAART 300
DB 241 VGLNDTMENTLAAVDRNEABISPSLTLYAACWENVPFINSPOPTFVGLIDLAIAART 300
QY 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYNHLGNNDGMNLSAPQFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYNHLGNNDGMNLSAPQFRSKEISKS 360
QY 361 NVVDDMVNSNALILYERGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVVDDMVNSNALILYERGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPITLIDLVLLAEISTRIEPAENEGKFSHPVATILSYLTKAPLVPGTPPVN 480
DB 421 EDSLLAAPITLIDLVLLAEISTRIEPAENEGKFSHPVATILSYLTKAPLVPGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 6
AAW79741
ID AAW79741 standard; protein; 510 AA.
XX
AC AAW79741;
XX
DT 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
XX
DE Soybean mutant myo-inositol 1-phosphate synthase.
XX
KM Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KM phytic acid.
XX
OS Glycine max; line LR33.
XX
PN WO9845448-A1.
XX
PD 15-OCT-1998.
XX
PF 07-APR-1998; 98WO-US006822.
XX
PR 08-APR-1997; 97US-00835751.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Hitz WD, Sebastian SA;

XX WPI: 1998-568353/48.
DR N-PSDB; AAW62443.
XX

PT Soybean plants containing altered myo-inositol 1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
PT phytic acid, etc.
XX

PS Example 5; Page 49-51; 63pp; English.

CC This is the amino acid sequence of a mutant soybean myo-inositol 1-
CC phosphate synthase (MI 1-PS) deduced from the coding region of an
CC isolated cDNA clone (see AAW62443). MI 1-PS is involved in glucose
CC metabolism to phytic acid, raffinose and stachyose. The MI 1-PS was
CC identified in soybean line LR33, a mutagenised line of low raffinose
CC saccharide phenotype. Sequencing revealed a single base change mutation
CC in the LR33 gene sequence that resulted in a K396N substitution in the
CC mutant protein compared to wild-type MI 1-PS (see AAW79740). The mutation
CC results in a seed phenotype of very low raffinose saccharide sugars, very
CC high sucrose and low phytic acid. The mutated nucleic acid is used to
CC alter the raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds, leading to useful soybean products,
CC e.g. a seed phytic acid content of less than 17 ug/g, a seed content of
CC raffinose and stachyose combined of less than 14.5 ug/g, and a seed
CC sucrose content greater than 200 ug/g. (Updated on 17-Oct-2003 to
CC standardise OS field)
XX

Sequence 510 AA;

Query Match 99.8%; Score 2627; DB 2; Length 510;
Best Local Similarity 99.8%; Pred. No. 3.1e-212;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MEIENFKVCEPNVKTETETISQVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
DB 1 MEIENFKVCEPNVKTETETISQVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
QY 61 KLGWNLVGMGNGNSTLTGCVIANREGISMAFKDKIQOANYGSLTQASAIRVGSFGQGE 120
DB 61 KLGWNLVGMGNGNSTLTGCVIANREGISMAFKDKIQOANYGSLTQASAIRVGSFGQGE 120
QY 121 IYAPFKSLIPMVPNDIVFGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPMVPNDIVFGGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDTIAANQEBRANNVIKGTQEQVOQIIKDIFKFKATKVDKVVVLTANTERYSNLV 240
DB 181 YDPDTIAANQEBRANNVIKGTQEQVOQIIKDIFKFKATKVDKVVVLTANTERYSNLV 240
QY 241 VGLNDTMENTLAAVDRNEABISPSLTLYAACWENVPFINSPOPTFVGLIDLAIAART 300
DB 241 VGLNDTMENTLAAVDRNEABISPSLTLYAACWENVPFINSPOPTFVGLIDLAIAART 300
QY 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYNHLGNNDGMNLSAPQFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYNHLGNNDGMNLSAPQFRSKEISKS 360
QY 361 NVVDDMVNSNALILYERGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVVDDMVNSNALILYERGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPITLIDLVLLAEISTRIEPAENEGKFSHPVATILSYLTKAPLVPGTPPVN 480
DB 421 EDSLLAAPITLIDLVLLAEISTRIEPAENEGKFSHPVATILSYLTKAPLVPGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 7
ADQ14495
ID ADQ14495 standard; protein; 510 AA.

XX	ADQ14495;
AC	
XX	23-SEP-2004 (first entry)
DT	
XX	
DE	Mutant soybean myo-inositol 1-phosphate synthase polypeptide #1.
XX	
KW	Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KW	raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
KW	mutant; mucin.
XX	
OS	Glycine max.
OS	Synthetic.
XX	
FT	Key Location/Qualifiers
FT	Misc-difference 396
XX	/note= "wild type Lys substituted by Asn"
XX	
PN	US2004128713-A1.
XX	
PD	01-JUL-2004.
XX	
PF	21-NOV-2003; 2003US-00718952.
XX	
PR	08-APR-1997; 97US-00835751.
PR	07-APR-1998; 98WO-US006822.
PR	26-APR-1999; 99US-00299315.
PR	11-MAR-2002; 2002US-00025003.
XX	
PA	(HITZ/) HITZ W D.
PA	(SEBA/) SEBASTIAN S A.
PA	(GRAC/) GRACE D J.
PA	(STREI/) STREIT L G.
XX	
PI	Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX	
DR	WPI; 2004-53135/51.
DR	N-PSDB; ADQ14494.
XX	
PT	New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT	useful for producing plants with decreased raffinose, stachyose, and
PT	phytic acid and increased sucrose, leading to valuable and useful soybean
PT	products.
XX	
PS	Claim 9; SEQ ID NO 6; 48pp; English.
XX	
CC	The invention relates to a nucleic acid fragment encoding a soybean myo-
CC	inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC	synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC	phosphate. The invention also relates to a chimeric gene operably linked
CC	to suitable regulatory sequences, where expression of the chimeric gene
CC	results in a decrease in expression of an endogenous or native gene
CC	encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC	comprising the chimeric gene, a method of making the soybean plant, a
CC	seed of the soybean plant, a soy protein product derived from the
CC	processing of soybean seeds, a method of making or producing a soy
CC	protein product and a method of using a soybean plant homozygous for at
CC	least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC	decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC	nucleic acid fragment and methods are useful for producing plants with
CC	decreased raffinose, stachyose and phytic acid content and increased
CC	sucrose and inorganic phosphate content, leading to valuable and useful
CC	soybean products. This sequence represents a mutant soybean myo-inositol
CC	1-phosphate synthase polypeptide of the invention.
XX	
90	Sequence 510 AA;

Query Match	99.8%;	Score 2627;	DB 8;	Length 510;
Best Local Similarity	99.8%;	Pred. No. 3.1e-212;		
Matches 509; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY 1 MFENFKCECPNVKLTETETIISVNVETTELVHNRNGYQMIIVPKSVKYEFTNNIHP 60

Db	1	MFLENPVECPNKKYETETISGVNMYETTELVEHNRNQTQWIVKPKSVKIEFKTNIHVP	60
Qy	61	KLGVMLVGMGNGNSLTLLTGVIANREGISWATKDKIQOANTFGSLTQASAIRVGSFOCEE	120
Db	61	KLGVMLVGMGNGNSLTLLTGVIANREGISWATKDKIQOANTFGSLTQASAIRVGSFOCEE	120
Qy	121	IYAPFSLILPMNPPDIVFGGMDISNMNLADAMARAKVPDIDLQKLRPYMESMLPLPGI	180
Db	121	IYAPFSLILPMNPPDIVFGGMDISNMNLADAMARAKVPDIDLQKLRPYMESMLPLPGI	180
Qy	181	YDPDFIAANOEEBRRANNVIKGTQOEQVOQIIKDKIKAFKEATKVDKVVLMVMTANTERYSNLV	240
Db	181	YDPDFIAANOEEBRRANNVIKGTQOEQVOQIIKDKIKAFKEATKVDKVVLMVMTANTERYSNLV	240
Qy	241	VGINDTMENLLAAVDSENEAISPSTLYAIACYMENVPPINGSPONTFVPGILDLAIARNT	300
Db	241	VGINDTMENLLAAVDSENEAISPSTLYAIACYMENVPPINGSPONTFVPGILDLAIARNT	300
Qy	301	LIGGDDPKSQGTQMKSVLVDFLVGAGIKRSTIVSTNHLGNDGNMLSAPOTFRSKEISKS	360
Db	301	LIGGDDPKSQGTQMKSVLVDFLVGAGIKRSTIVSTNHLGNDGNMLSAPOTFRSKEISKS	360
Qy	361	NVVDMDMNSNAIIYYEGEHPDHVVVIXKYVYVGDSSKRAMEDEYTSSEIFMGKSTIVLHNTC	420
Db	361	NVVDMDMNSNAIIYYEGEHPDHVVVIXKYVYVGDSSKRAMEDEYTSSEIFMGKSTIVLHNTC	420
Qy	421	EDSLAAPRIILDLVLAELSTRLEFKAENEGKFSFHPVATILSYLTKAPLVPPGTPVYN	480
Db	421	EDSLAAPRIILDLVLAELSTRLEFKAENEGKFSFHPVATILSYLTKAPLVPPGTPVYN	480
Qy	481	ALSKORAMLENIRACVGLAPENNMILEYK 510	
Db	481	ALSKORAMLENIRACVGLAPENNMILEYK 510	
RESULT 8			
ADs81998	ID	ADs81998 standard; protein; 510 AA.	
XX	AC	ADs81998;	
XX	DT	18-NOV-2004 (first entry)	
DE	XX	Soybean myo-inositol 1-phosphate synthase mutant #1.	
XX	XX		
KM	XX	Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; mutein;	
KM	XX	phytic acid; raffinose; stachyose; sucrose; inorganic phosphate;	
KM	XX	flutulence; mutant.	
OS	XX	Glycine max; line LR33.	
OS	XX	Synthetic.	
XX	XX		
XX	XX		
FT	Key	Location/Qualifiers	
FT	Misc-difference 396	/note= "wild-type Lys substituted by Arg"	
XX	PN	US2003074685-A1.	
XX	PD	17-APR-2003.	
XX	PF	11-MAR-2002; 2002US-00025003.	
XX	PR	08-APR-1997; 97US-00835751.	
XX	PR	07-APR-1998; 98WO-US006622.	
XX	PA	(HITZ/) HITZ W D.	
XX	PA	(SEBA/) SEBASTIAN S A.	
XX	PI	Hitz WD, Sebastian SA;	
DR	XX	WPI; 2004-639957/62.	
DR	XX	N-PSDB; ADs81997.	
XX	XX		

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

PS Claim 9; SEQ ID NO 6; 34bp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homologous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate), where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence represents a mutant myo-
CC inositol 1-phosphate synthase.

CC Sequence 510 AA;

Query Match 99.8%; Score 2627; DB 8; Length 510;

Best Local Similarity 99.8%; Pred. No. 3.1e-212;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEIENFKVCEPNVKTETETISQVYNYETTELHVENNGYOMIVKRSKYTEFTNINHP 60
DB 1 MEIENFKVCEPNVKTETETISQVYNYETTELHVENNGYOMIVKRSKYTEFTNINHP 60
QY 61 KGVNVLVGGNGNSTLTGSGVIANREGISMAFKDKIQOANYSGLTQASAIRVSGQGE 120
DB 61 KGVNVLVGGNGNSTLTGSGVIANREGISMAFKDKIQOANYSGLTQASAIRVSGQGE 120
QY 121 IYAPFKSLPMVNPDDIVEGWDISNMNLADAMARAKVDFDILQOLRPMESMLPLPGI 180
DB 121 IYAPFKSLPMVNPDDIVEGWDISNMNLADAMARAKVDFDILQOLRPMESMLPLPGI 180
QY 181 YPPDPTIAANQERANNVIGTKQEQVQIILKIDKAFKATKVDKVVLTANTERYSNIV 240
DB 181 YPPDPTIAANQERANNVIGTKQEQVQIILKIDKAFKATKVDKVVLTANTERYSNIV 240
QY 241 YPDPPIAANQERANNVIGTKQEQVQIILKIDKAFKATKVDKVVLTANTERYSNIV 240
DB 241 YPDPPIAANQERANNVIGTKQEQVQIILKIDKAFKATKVDKVVLTANTERYSNIV 240
QY 241 VGLNPTMENTLAAVDNENAEISPSLTAAACWENVPFINGSPONTFVGLDILAIARNT 300
DB 241 VGLNPTMENTLAAVDNENAEISPSLTAAACWENVPFINGSPONTFVGLDILAIARNT 300
QY 301 ILGGDDPKSGQTKMSVLVDFLVGAGIKRPSIVSYNHLGNNDGMNLASQTRSKSISKS 360
DB 301 ILGGDDPKSGQTKMSVLVDFLVGAGIKRPSIVSYNHLGNNDGMNLASQTRSKSISKS 360
QY 361 NVVDDMVNSNALLYEGEHPDHVVVVKYPPYGDSCRANDTTSSEIFMGSKSTIVLANTC 420
DB 361 NVVDDMVNSNALLYEGEHPDHVVVVKYPPYGDSCRANDTTSSEIFMGSKSTIVLANTC 420
QY 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKPHSHFPAATILSYTKRPLVPPGTPVYN 480
DB 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKPHSHFPAATILSYTKRPLVPPGTPVYN 480

DB 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKPHSHFPAATILSYTKRPLVPPGTPVYN 480
QY 481 ALSKQPMLENTMRACVGLAPENNMLEYK 510
DB 481 ALSKQPMLENTMRACVGLAPENNMLEYK 510

RESULT 9
ADQ14503
ID ADQ14503 standard; protein; 510 AA.

AC ADQ14503;

DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #4.

XX Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;

KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;

KW mutant; mutein.

XX Glycine max.

XX Synthetic.

PN US2004128713-A1.

PD 01-JUL-2004.

PF 21-NOV-2003; 2003US-00718952.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.

PR 26-APR-1999; 99US-00299315.

PR 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.

PA (STRE/) STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

XX WPI, 2004-533135/51.

DR N-PSDB; ADQ14502.

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,

PT useful for producing plants with decreased raffinose, stachyose, and

PT phytic acid and increased sucrose, leading to valuable and useful soybean

PT products.

XX Example 8; SEQ ID NO 14; 48bp; English.

PS The invention relates to a nucleic acid fragment encoding a soybean myo-

XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate

CC synthase having a decreased capacity for the synthesis of myo-inositol 1-

CC phosphate. The invention also relates to a chimeric gene operably linked

CC to suitable regulatory sequences, where expression of the chimeric gene

CC results in a decrease in expression of an endogenous or native gene

CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant

CC comprising the chimeric gene, a method of making the soybean plant, a

CC seed of the soybean plant, a soy protein product derived from the

CC processing of soybean seeds, a method of making or producing a soy

CC protein product and a method of using a soybean plant homozygous for at

CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having

CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The

CC nucleic acid fragment and methods are useful for producing plants with

CC decreased raffinose, stachyose and phytic acid content and increased

CC sucrose and inorganic phosphate content, leading to valuable and useful

CC soybean products. This sequence represents a mutant soybean myo-inositol

CC 1-phosphate synthase polypeptide of the invention.

XX Sequence 510 AA;

Query Match	98.8%	Score 2600;	DB 8;	Length 510;
Best Local Similarity	98.6%	Pred. No. 5.9e-210;		
Matches 503; Conservative	4;	Mismatches 3;	Indels 0;	Gaps 0

Qy	1	MFLENKRVCEPNKKTETETELQSYVNNVETTELVAHNBNCTYQWTVYKPKSYKFEFKNIHV	60
Db	1	MFLENKRVESPNNKYETETELQSYVNNVETTELVAHNBNCTYQWTVYKPKSYVYQFKNIHV	60
Qy	61	KLGVMVLVGMCGNNGSLTGTGVIANREGISWATYKDKIQANVFGSLTQASAIWVSGFOGE	12
Db	61	KLGVMVLVGMCGNNGSLTGTGVIANREGISWATYKDKIQANVFGSLTQASAIWVSGFOGE	12
Qy	121	IYAPFKSLLPMVNPDDIVFEGWDISNMNLADAMARAKVFDIDLQKLRPYMESMLPLFGI	18
Db	121	IYAPFKSLLPMVNPDDIVFEGWDISNMNLADAMARAKVFDIDLQKLRPYMESMPLFGI	18
Qy	181	YDPDFILAAQOEERANNVIKGTKEQVQOQIIKIKAKKATKYDKYVLTMTATERYSLV	24
Db	181	YDPDFILAAQOEERANNVIKGTKEQVQOQIIKIKAKKATKYDKYVLTMTATERYSLV	24
Qy	241	VGNDPMENLLAVDNRNEAISPSTLYALACWENVPFNGSPONTFVPGLLDLIARNT	30
Db	241	VGNDPMENLLAVDNRNEAISPSTLYALACWENVPFNGSPONTFVPGLLDLIARNT	30
Qy	301	LIGDDPFKSGQTCKMSVLVDPLGAIKIKETSIYSVYHNLGNDGMLSAQTFRSKEISKS	36
Db	301	LIGDDPFKSGQTCKMSVLVDPLGAIKIKETSIYSVYHNLGNDGMLSAQTFRSKEISKS	36
Qy	361	NVYDDMWNSSALLIYEPGEHPDHVVYIKYVPYGDGSRADERTYSEIFMGKSTIVLANTC	42
Db	361	NVYDDMWNSSALLIYEPGEHPDHVVYIKYVPYGDGSRADERTYSEIFMGKSTIVLANTC	42

[illegible]

RESULT 10
ADQ14505
ID ADQ14505 standard; protein; 510 AA

AC ADQ14505 ;

DT 23-SEP-2004 (first entry)

DE Wild type soybean myo-inositol 1-phosphate synthase polypeptide #2.

KM Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KM raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme

Glycine max.

PN US2004128713-A1.

PD 01-JUL-2004.

PF 21-NOV-2003; 2003US-00718952.

PR 08-APR-1997; 97US-00835751.

PR 26-APR-1999; 99US-00299315.

XX

PA (SEBA/) SEBASTIAN S A.

PA (STRE/) STREIT L G.

PI Hitz WD, Sebastian SA

XX WPI, 2004-533135/51.
DR N-PSDB; ADQ14504.
DR
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

Claim 3; SEQ ID NO 16; 48pp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and pyruvic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a wild type soybean myo-
CC inositol 1-phosphate synthase polypeptide of the invention.

SQ Sequence 510 AA;

Query Match	98.8%	Score 2600;	DB 8;	Length 510;
Best Local Similarity	98.6%	Pred. No. 5.9e-210;		
Matches 503; Conservative	4;	Mismatches 3;	Indels 0;	Gaps 0

Qy	1	MFLENKVECPNKYKTEETISQSVNYEYETTEL,VHEHNRKNTQOMIKPSPSVYEFKNTNHP	60
Db	1	MFLENKVESPNVKYTEIEBISQSVNYEETTEL,VHEHNRKNTQOMIKPSPSVYQEFKNTNHP	60
Qy	61	KLGVM,VLGNGNGNSTLTGTGVIANREGISMA7TKDIOQANYFGSL7QASAIRVGSFOGEE	120
Db	61	KLGVM,VLGNGNGNSTLTGTGVIANREGISMA7TKDIOQANYFGSL7QASAIRVGSFOGEE	120
Qy	121	IYAPEFSLL,PMVNPDDIVFGMDISNNMLAAMARAKYFDIDLOKOLRPWESMLPLPGI	180
Db	121	IYAPFSLL,PMVNPDDIVFGMDISNNMLADAMARAKYFDIDLOKOLRPWESMLPLPGI	180
Qy	181	YDPDFIAANQEEBRRANVIKGTQOEVOQI1KD1KAFKATKVDKVVYLM7ANTERYSNLV	240
Db	181	YDPDFIAANQEEBRRANVIKGTQOEVOQI1KD1KAFKATKVDKVVYLM7ANTERYSNLV	240
Qy	241	VGLNDMMENLIAVUBRNEAETSPSLTYAIACMEVVRPINGSPONFEPVGLIDATAARNT	300
Db	241	VGLNDMMENLIAVUBRNEAETSPSLTYAIACMEVVRPINGSPONFEPVGLIDATAARNT	300
Qy	301	LIGGDDFKSGQTKMSVLVDPLVGAGIKPRTSIVSYNHLGNNDGNMLAPOTFRSKEISKS	360
Db	301	LIGGDDFKSGQTKMSVLVDPLVGAGIKPRTSIVSYNHLGNNDGNMLAPOTFRSKEISKS	360
Qy	361	NAVDDMVNSNALLYEPGEHPDHVVYIKVVPVVGSKRAMDEYTSIEMFGGNNITVLHNTC	420
Db	361	NAVDDMVNSNALLYEPGEHPDHVVYIKVVPVVGSKRAMDEYTSIEMFGGNNITVLHNTC	420
Qy	421	EDSLLAAPITLLDVLIAELSTR1EFKANEKGKFSFHPVATILISYLTAKLAPVPGTPVAN	480
Db	421	EDSLLAAPITLLDVLIAELSTR1EFKANEKGKFSFHPVATILISYLTAKLAPVPGTPVAN	480
Qy	481	ALSKORAM,ENIMRACVGLAEPNNMILSYK	510
Db	481	ALSKORAM,ENIMRACVGLAEPNNMILSYK	510

	RESULT 11
ADSB82004	ID ADSB82004 standard; protein; 510 AA.
XX	
XX AC	ADSB82004;
XX XX	
XX	18-NOV-2004 (first entry)
DE	Soybean myo-inositol 1-phosphate synthase wild-type 4.
XX	
KW	Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid; raffinose; starchose; sucrose; inorganic phosphate; flatulence.
OS	Glycine max; line 2901BJP03.
XX	
PN	US2003074685-A1.
PD	
PD	17-APR-2003.
PF	
PF	11-MAR-2002; 2002US-00025003.
PR	
PR	08-APR-1997; 97US-00835751.
PR	07-APR-1998; 98WO-US006822.
XX	
PA	(HITZ/) HITZ W D.
PA	(SEBA/) SEBASTIAN S A.
PI	Hitz WD, Sebastian SA;
DR	
DR	WPI: 2004-639857/62.
DR	N-P5DB; ADSB82003.
PT	
PT	Novel isolated nucleic acid fragment encoding soybean myo-inositol 1- phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
XX	
XX	Example 8; SEQ ID NO 14; 34pp; English.
XX	
CC	The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1- phosphate synthase having decreasing capacity for the synthesis for myo- inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytia acid content of less than 17 micromol/g, a seed content of raffinose plus starchose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homologous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with for heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a wild-type myo- inositol 1-phosphate synthase.
XX	
XX	Sequence 510 AA;
SO	
Query Match	98.8%; Score 2600; DB 8; Length 510;
Best Local Similarity	98.6%; Pred No. 5; 9e-210;

	Matches	503; Conservative	4; Mismatches	3; Indels	0; Gaps	0; PT
QY		1	MFENFKVECPNVKXTTETIQSYVNYETTEL	VHNRNGTYQWIKPKSVKKEFKTNTHVP	60	
Db		1	MFENFKVESPNVKXTTETIQSYVNYETTEL	VHNRNGTYQWIKPKSVNYQFNTNTHVP	60	
QY		61	KLGVMLVWGGNGNSTLTGVIANREGISWATKDKIQOANTFGSLTQASAIRVSGFOEE	120		
Db		61	KLGVMLVWGGNGNSTLTGVIANREGISWATKDKIQOANTFGSLTQASAIRVSGFOEE	120		
QY		121	IYAPFKSLIPVNPDDIVIFGGMDISNMNLADAMARAKVFDIDLKOLRPYMESMLPLPGI	180		
Db		121	IYAPFKSLIPVNPDDIVIFGGMDISNMNLADAMARAKVFDIDLKOLRPYMESVPLPGI	180		
QY		181	YDPPFIANAQBERANNVIKGTKEQVOQOIIDIKAFKEATKVDKVVVLMTANTERYSNLV	240		
Db		181	YDPPFIANAQBERANNVIKGTKEQVOQOIIDIKAFKEATKVDKVVVLMTANTERYSNLV	240		
QY		241	VGLNDTMENLLAAVDRNEAEISPSSTLYAIACWMENVPPINGSPONTFVPGILIDIALANT	300		
Db		241	VGLNDTMENLLAAVDRNEAEISPSSTLYAIACWMENVPPINGSPONTFVPGILIDIALANT	300		
QY		301	LIGGDDFFSGGTTKKSVLYVDFVAGTKPTISIVSYNHLGNDGNLSPQFRSKAISKS	360		
Db		301	LIGGDDFFSGGTTKKSVLYVDFVAGTKPTISIVSYNHLGNDGNLSPQFRSKAISKS	360		
QY		361	NVVDMDVNSNAILLYEPGEHPDHVVVIKVPYPVGSKRAMDEYTSIEIFMGKSTIVLHNTC	420		
Db		361	NVVDMDVNSNAILLYEPGEHPDHVVVIKVPYPVGSKRAMDEYTSIEIFMGKNTIVLHNTC	420		
QY		421	EDSLIAPRIIDLVLLAELSTRIEFKKAENEGKFFHSFHPVATILSYLTAKPLVPPTPVVN	480		
Db		421	EDSLIAPRIIDLVLLAELSTRIOFKKAENEGKFFHSFHPVATILSYLTAKPLVPPTPVVN	480		
QY		481	ALSKORAMLENIMRACVGLAENNMILLEKX	510		
Db		481	ALSKORAMLENIMRACVGLAENNMILLEKX	510		
RESULT 12						
ADSB2006	ID	ADSB2006	standard; protein; 510 AA.			
XX	AC	ADSB2006;				
XX	DT	18-NOV-2004	(first entry)			
XX	DE	Soybean myo-inositol 1-phosphate synthase wild-type 2.				
XX	KW	Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid; raffinose; stachyose; sucrose; inorganic phosphate; flatulence.				
XX	OS	Glycine max; cultivar Wye.				
XX	PN	US2003074685-A1.				
XX	PD	17-APR-2003.				
XX	PF	11-MAR-2002; 2002US-00025003.				
XX	PR	08-APR-1997; 97US-00835751.				
XX	PR	07-APR-1998; 98WO-US006822.				
XX	PA	(HITZ/) HITZ W D.				
XX	PA	(SEBA/) SEBASTIAN S A.				
XX	PI	Hitz WD, Sebastian SA;				
XX	DR	WPI; 2004-639957/62.				
XX	DR	N-PSDB; ADSB2005.				
PT	Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose,					

PT phytic acid and inorganic phosphate content of soybean seeds.
XX
PS Claim 3; SEQ ID NO 16; 34pp; English.
XX
CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with a
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flautence
CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.
SQ Sequence 510 AA:

Query Match 98.8%; Score 2600; DB 8; Length 510;
Best Local Similarity 98.6%; Pred. No. 5,9e-210;
Matches 503; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MFENFKVCEPNVKTETETIGSVNYETTELVEHNRNGTYQWIVPKSVKEFKTNHVP 60
DB 1 MFENFKVCEPNVKTETETIGSVNYETTELVEHNRNGTYQWIVPKSVKEFKTNHVP 60
QY 61 KLGWLVGNGNGSTLTGVIANREGISWATKDKIQQANNYGSITQASAIRVGSFOGEE 120
DB 61 KLGWLVGNGNGSTLTGVIANREGISWATKDKIQQANNYGSITQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPWVNPDDIVFGWMDISNMNLADAMAKVYEDIDLOQLRPYMSMLPLPGI 180
DB 121 IYAPFKSLIPWVNPDDIVFGWMDISNMNLADAMAKVYEDIDLOQLRPYMSMLPLPGI 180
QY 181 YDPDFIAANQEEERANNVTKGTQEOVOQIIKQIKAFKEATKYDKVYVMTANTERYSNLY 240
DB 181 YDPDFIAANQEEERANNVTKGTQEOVOQIIKQIKAFKEATKYDKVYVMTANTERYSNLY 240
QY 241 VGLNMTMENVLAAVNRNEAISPSTLYAACWENVPFINSPOMTFVPGILDIAIART 300
DB 241 VGLNMTMENVLAAVNRNEAISPSTLYAACWENVPFINSPOMTFVPGILDIAIART 300
QY 301 LIIGDDFKSGQTKMSVLVDFVGAGIKPTSTIVSYNHLGNNNGMLSAPQTRSRKEISKS 360
DB 301 LIIGDDFKSGQTKMSVLVDFVGAGIKPTSTIVSYNHLGNNNGMLSAPQTRSRKEISKS 360
QY 361 NVVDDMVNSNALIYEPGEHPDHVVVTKYVPYGDSCRADDEYTSSEIFMGKSTIYLAHTC 420
DB 361 NVVDDMVNSNALIYEPGEHPDHVVVTKYVPYGDSCRADDEYTSSEIFMGKSTIYLAHTC 420
QY 421 EDSLLAAPITLVLVLAELSTRIFPKAENEGKFSHPATLSTLYTAKAIVLPQGPVNV 480
DB 421 EDSLLAAPITLVLVLAELSTRIFPKAENEGKFSHPATLSTLYTAKAIVLPQGPVNV 480
SQ

QY 481 ALSKQRAWLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKQRAWLENIMRACVGLAPENNMILEYK 510

RESULT 13
ADQ14501
ID ADQ14501 standard; protein, 510 AA.
XX
AC ADQ14501;
XX
DI 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #3.
DE Soybean: myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KM raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
KM mutant; mutcin.
XX
OS Glycine max.
OS Synthetic.
XX
PN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US0006822.
PR 26-APR-1999; 99US-00293315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
DR WPI, 2004-533135/51.
DR N-PSDB; ADQ14500.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
PS Claim 9; SEQ ID NO 12; 48pp; English.
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.
SQ Sequence 510 AA:

Query Match 98.5%; Score 2593; DB 8; Length 510;
Best Local Similarity 98.4%; Pred. No. 2.3e-209;

Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MEIENFKCEPVKKTETETISQSVNYETTELHNRNGTYQIVPKSVKYEKTNINHP 60
 DB 1 MEIENFKVESPVKKTETETISQSVNYETTELHNRNGTYQIVPKSVNYQFKTNTHVP 60
 QY 61 KLGWLVGVGNGNSTLTGCVIANREGISMAWKDKIQOANYGSLTQASAIRVGSFGQEE 120
 DB 61 KLGWLVGVGNGNSTLTGCVIANREDISMAWKDKIQOANYGSLTQASAIRVGSFGQEE 120
 QY 121 IYAPFKSLPVMNPPDIYFGGWDISNMNLADAMARAKVPDIDLOQLARPYMESMLPLPGI 180
 DB 121 IYAPFKSLPVMNPPDIYFGGWDISNMNLADAMARAKVPDIDLOQLARPYMESMLPLPGI 180
 QY 181 YDPDFIAANOBERANNAVIGTKQEOVOQI IKDIFKFEKATKYDKVVLMTANTERYSNLV 240
 DB 181 YDPDFIAANOBERANNAVIGTKQEOVOQI IKDIFKFEKATKYDKVVLMTANTERYSNLV 240
 QY 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWMEVNFINGSPONTFVGLIDLAIARNT 300
 DB 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWMEVNFINGSPONTFVGLIDLAIARNT 300
 QY 301 LIGGDDPKSGQTKMSVLVDPLVAGAGIKPTSTIVSYNHLGNNDGMNLAPQTRSKSISKS 360
 DB 301 LIGGDDPKSGQTKMSVLVDPLVAGAGIKPTSTIVSYNHLGNNDGMNLAPQTRSKSISKS 360
 QY 361 NVVDDMVNSNALIYERGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
 DB 361 NVVDDMVNSNALIYERGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
 QY 421 EDSLLAAPTILDLVLAELSTRIEPKAENEGKPHSPVATILSYLTRKPLVPPGTPVYN 480
 DB 421 EDSLLAAPTILDLVLAELSTRIEPKAENEGKPHSPVATILSYLTRKPLVPPGTPVYN 480
 QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
 DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 14
 ADS82002
 ID ADS82002 standard; protein; 510 AA.
 XX
 AC ADS82002;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Soybean myo-inositol 1-phosphate synthase mutant #2.
 XX
 KM Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; mutein;
 KM phytic acid; raffinose; stachyose; sucrose; inorganic phosphate;
 KM flatulence; mutant.
 XX
 OS Glycine max; line 29010CP01.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 87
 FT /note= "wild-type Gly substituted by Asp"
 XX
 PN US2003074685-A1.
 PD 17-APR-2003.
 PF 11-MAR-2002; 2002US-00025003.
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 XX
 PI Hitz WD, Sebastian SA;

XX
 DR MPI: 2004-639957/62.
 DR N-PSDB; ADS82001.
 XX
 PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
 PT
 PS Claim 9; SEQ ID NO 12; 34pp; English.
 XX
 CC The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence CC when consumed by humans. The present sequence represents a mutant myo-inositol 1-phosphate synthase.
 CC
 XX

SO Sequence 510 AA;

Query Match 98.5%; Score 2593; DB 8; Length 510;
 Best Local Similarity 98.4%; Pred. No. 2.3e-209;
 Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MEIENFKCEPVKKTETETISQSVNYETTELHNRNGTYQIVPKSVKYEKTNINHP 60
 DB 1 MEIENFKVESPVKKTETETISQSVNYETTELHNRNGTYQIVPKSVNYQFKTNTHVP 60
 QY 61 KLGWLVGVGNGNSTLTGCVIANREGISMAWKDKIQOANYGSLTQASAIRVGSFGQEE 120
 DB 61 KLGWLVGVGNGNSTLTGCVIANREDISMAWKDKIQOANYGSLTQASAIRVGSFGQEE 120
 QY 121 IYAPFKSLPVMNPPDIYFGGWDISNMNLADAMARAKVPDIDLOQLARPYMESMLPLPGI 180
 DB 121 IYAPFKSLPVMNPPDIYFGGWDISNMNLADAMARAKVPDIDLOQLARPYMESMLPLPGI 180
 QY 181 YDPDFIAANOBERANNAVIGTKQEOVOQI IKDIFKFEKATKYDKVVLMTANTERYSNLV 240
 DB 181 YDPDFIAANOBERANNAVIGTKQEOVOQI IKDIFKFEKATKYDKVVLMTANTERYSNLV 240
 QY 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWMEVNFINGSPONTFVGLIDLAIARNT 300
 DB 241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWMEVNFINGSPONTFVGLIDLAIARNT 300
 QY 301 LIGGDDPKSGQTKMSVLVDPLVAGAGIKPTSTIVSYNHLGNNDGMNLAPQTRSKSISKS 360
 DB 301 LIGGDDPKSGQTKMSVLVDPLVAGAGIKPTSTIVSYNHLGNNDGMNLAPQTRSKSISKS 360
 QY 361 NVVDDMVNSNALIYERGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420

Db 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKAPLVPGTTPVNV 480
Db 421 EDSLAAPIIIDLVLAELSTRIOFKAENEGKFSFHPVATILSYLTAKAPLVPGTTPVNV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510

RESULT 15

AAV24477
ID AAV24477 standard; protein; 536 AA.

XX AC AAV24477;

DT 24-SEP-1999 (first entry)

XX Nicotiana paniculata INPS protein.

XX Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;

KW water stress; resistance.

XX Nicotiana paniculata.

XX Key Location/Qualifiers

FT Misc-difference 511

FT //label= unknown

FT /note= "encoded by the stop codon TGA"

PN JP1187879-A.

XX 13-JUL-1999.

XX 26-DEC-1997; 97JP-00359773.

XX 26-DEC-1997; 97JP-00359773.

XX (NIBS) JAPAN TOBACCO INC.

XX WPI; 1999-451546/38.

DR N-PSDB; AAX90402.

PT New INPS gene derived from Nicotiana genus plant - useful for conferring
PT resistance to water stress to plants.

XX Claim 2; Page 6-8; Bpp; Japanese.

CC The present sequence is the Nicotiana paniculata inositol monophosphate
CC synthase (INPS), designated NpINPS1. INPS can be used to confer water
CC stress resistance to a plant

XX Sequence 536 AA;

Query Match 93.9%; Score 2472; DB 2; Length 536;

Best Local Similarity 92.2%; Pred. No. 3.8e-199;

Matches 470; Conservative 24; Mismatches 16; Indels 0; Gaps 0;

Qy 1 MEIENFKVECPNVKYTEIETIQSVYNYETTELVHENRNGTYOMIVKPKSVKYEFTNIHVP 60
Db 1 MEIENFKVESPVVKYTESRHSVYDQTELVHDEKNGTYQWTVKPKTVKYEFTDVHVP 60
Qy 61 KLGVMVLVGKGGNGSLTGGVIANREGISWATKDKVQANNFGSLTQASATIRVGSFGQGE 120
Db 61 KLGVMVLVGKGGNGSLTGGVIANREGISWATKDKVQANNFGSLTQASATIRVGSFGNGB 120
Qy 121 IYAPFKSLPLMVPNPDIVFGWDISNMNLADAMARAKVFDIDLOKQLRPYMESMLPLPGI 180
Db 121 IYAPFKSLPLMVPNPDIVFGWDISNMNLADAMARAKVFDIDLOKQLRPYMESMLPLPGI 180
Qy 181 YDPDFIAANQERANNVIKGTQBOVQOIIKDIKAFKEATKVDKVVVLTANTERYSNLV 240
Db 181 YDPDFIAANQERANNVIKGTQBOVQOIIKDIKAFKEATKVDKVVVLTANTERYSNLV 240

Db 181 YDPDFIAANQSRANNVIKGTKEQIDQIIKDIREFKEKNKVDKVVVLTANTERYSNV 240
Qy 241 VGLNDTMENLAAVDRNEAEISPTLYAIACWENVPFINGSPONTFVPGILDLAIARNT 300
Db 241 VGLNDTMENLFAVDNRNABEISPTLYAIACILENVPFINGSPONTFVPGILDLAIKNT 300
Qy 301 LIIGDDFKSGQTKKKSIVLVFLVGAQIKPTSIIVSYNHLGNNDGMNLAPOTFRSKEISKS 360
Db 301 LIIGDDFKSGQTKKKSIVLVFLVGAQIKPTSIIVSYNHLGNNDGMNLAPOTFRSKEISKS 360
Qy 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKAPLVPGTTPVNV 480
Db 421 EDSLAAPIIIDLVLAELSTRIOFKAENEGKFSFHPVATILSYLTAKAPLVPGTTPVNV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510

Search completed: June 7, 2005, 16:43:36
Job time : 99.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:16:29 ; Search time 24.5 Seconds
(without alignments)
2002.880 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632

Sequence: 1 MFIEHFKECPNVKYTEIRI.....NIMRACVGLAPENNMLEYK 510

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*

- 1: PIR1:*
- 2: PIR2:*
- 3: PIR3:*
- 4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2381	90.5	512	2 T12438	inositol-3-phospha
2	2378	90.3	510	2 D84610	probable myo-inosi
3	2369	90.0	510	2 T50021	inositol-3-phospha
4	2363	89.8	509	2 T08436	inositol-3-phospha
5	2352.5	89.4	511	2 T05017	inositol-3-phospha
6	2349	89.2	510	2 S60302	inositol-3-phospha
7	2344	89.1	510	2 T01647	inositol-3-phospha
8	2331.5	88.6	511	2 T10964	inositol-3-phospha
9	2328	88.4	510	2 T04399	inositol-3-phospha
10	2314.5	87.9	507	2 S52648	inositol-3-phospha
11	1386.5	52.8	525	2 T18569	inositol-3-phospha
12	1382.5	52.5	555	2 A30902	inositol-3-phospha
13	1381.5	52.5	520	2 S45452	inositol-3-phospha
14	1308.5	49.7	430	2 T46317	inositol-3-phospha
15	296	11.2	388	2 T36191	probable myo-inosi
16	285.5	10.8	392	2 A69474	myo-inositol-1-pho
17	242	9.2	417	2 T34930	probable secreted
18	214	8.1	382	2 H72255	myo-inositol-1-pho
19	212	8.1	425	2 E71039	hypothetical prote
20	198	7.5	386	2 B75175	myo-inositol-1-pho
21	195	7.4	364	2 G70451	conserved hypothet
22	163	6.2	360	2 B90239	conserved hypothet
23	158	6.0	392	2 F72632	probable myo-inosi
24	145.5	5.5	368	2 A65014	conserved hypothet
25	142	5.4	369	2 S72835	hypothetical prote
26	133	5.1	360	2 T36586	hypothetical prote
27	128.5	4.9	367	2 F70912	hypothetical prote
28	125.5	4.8	367	2 G95871	conserved hypothet
29	119.5	4.5	873	2 T16282	hypothetical prote

30	118	4.5	739	2 B86816	hypothetical prote
31	116.5	4.4	344	2 B90176	conserved hypothet
32	116	4.4	535	2 A80983	periplasmic dipepc
33	115.5	4.4	731	2 D95235	penicillin-binding
34	115.5	4.4	1993	2 AF1450	probable peptidogl
35	115	4.4	2490	1 A54971	protein-tyrosine-p
36	114.5	4.4	867	2 B86815	ClbB protein (limo
37	114.5	4.4	900	2 S46101	hypothetical prote
38	113	4.3	351	2 D82930	SRP family of GTP-
39	113	4.3	739	2 T51701	phosphoribosylform
40	113	4.3	802	2 C90326	hypothetical prote
41	113	4.3	4936	2 A42515	hypothetical prote
42	112.5	4.3	430	2 T04668	phosphoserine tran
43	112.5	4.3	867	2 H90524	preprotein translo
44	111.5	4.2	731	2 B98099	peptidoglycan glyc
45	111.5	4.2	776	2 S59790	hypothetical prote

ALIGNMENTS

Query Match	90.5%	Score 2381	DB 2	Length 512
Best Local Similarity	88.7%	Pred. No. 4.6e-153		
Matches	454	Conservative	31	Mismatches 25; Indels 2; Gaps 1
QY	1	MFIEHFKECPNVKYTEIRIISVNYETTELVEHKN--GTYQIVKRSVKYERKTH 58		
DB	1	MFIEBFKESPVVKYTEIRIESVNYDTTELVEHNRKQAGYQWIVKPTVGYHRTDTR 60		
QY	59	VFKLGVMLVGMGNGNSTLTGIVINRSGISWATKDKIQQANYFQSLTQASAIRVSGFQG 118		
DB	61	VFKLGVMLVGMGNGNSTLTGIVINRSGISWATKDKIQQANYFQSLTQASIRVSGFNG 120		
QY	119	BEIVAFPKSLFPMVNPDDIVFGMDISNMNLADAMARAKVPDIDLOKOLRPYMSMLPLP 178		
DB	121	BEIVAFPKSLFPMVNPDDIVFGMDISNMNLADAMARAKVPDIDLOKOLRPYMSMLPLP 180		
QY	179	GIYDPDFIAANDPEFANRVYIKTKQEQVQIITKIKAFREATKVDKVVVLTANTERRYSN 238		
DB	181	GIYDPDFIAANDPEFANRVYIKTKQEQVQIITKIKAFREATKVDKVVVLTANTERRYSN 240		
QY	239	LVVGGINDMENLAAVDNENAEISPTLYATAQWENPFIINGSPONTFVPELIDLAIR 298		
DB	241	VVVGINDMENLAAVDNENAEISPTLYATAQWENPFIINGSPONTFVPELIDLAIR 300		
QY	299	NTLIGDDPFKSGQTQKMSVLVDVFLGAGIKPSTISVYHNLGNDGMSAPQTFPSKES 358		
DB	301	NSLIGDDPFKSGQTQKMSVLVDVFLGAGIKPSTISVYHNLGNDGMSAPQTFPSKES 360		
QY	359	KSNVVDMAVNSNAIYYEPGEHPDHVVVTKYPPYVDSRPANDYTSSEIFMGSKSTIVLHN 418		
DB	361	KSNVVDMAVNSNAIYYEPGEHPDHVVVTKYPPYVDSRPANDYTSSEIFMGSKSTIVLHN 420		
QY	419	TCEDSLAAPITLIDVLAELSTRIEFKAENGGKSHSPAVATISYLTAKPLVPRGTPV 478		

Db 421 TCEDSLAAPILLDLVLAELSTRIOQLKAEEDKHFHPVATILSYLTAKAPLVPGTPV 480
Qy 479 VNALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 VNALSKORAMLENIMRACVGLAPENNMILEYK 512

RESULT 2
D84610
probable myo-inositol 1-phosphate synthase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Aug-2004
C:Accession: D84610
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
Nature 402, 761-766, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84610
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-510 <STO>
A:Cross-references: UNIPROT:Q38662; GB:AE002093; NID:g4567202; PIDN:AAD33618.1; GSPDB:GN
C:Genetics:
A:Gene: At2g22240
A:Map position: 2
C:Superfamily: Myo-inositol-1-phosphate synthase

Query Match 90.3%; Score 2378; DB 2; Length 510;
Best Local Similarity 87.8%; Pred. No. 7.3e-153;
Matches 448; Conservative 40; Mismatches 22; Indels 0; Gaps 0;

Qy 1 MEIENFKVCEPNVKKYTEIETSIQSVNYETTELVHENRNGTYQWIVPKSKYKEFKTNHVP 60
Db 1 MFIESFKVESPNVKKYTEIENISVVDYETTELVHENRNGTYQWIVPKSKYKEFKTDTRVP 60
Qy 61 KLGVWLVGWGNNGSTLTGVIANREGISMAWKIKIQOANVFGLTQASIRVGSFOGEE 120
Db 61 KLGVWLVGWGNNGSTLTGVIANREGISMAWKIKQOANVFGLTQASSIRVGSNGEE 120
Qy 121 IYAPFKSLIPWNPDDIYFGGWDISNNMADAMARAKVFDIDLOKQLRPYMESMLPLPGI 180
Db 121 IYAPFKSLIPWNPDDIYFGGWDISNNMADAMARAKVFDIDLOKQLRPYMENMPLPGI 180
Qy 121 IYAPFKSLIPWNPDDIYFGGWDISNNMADAMARAKVFDIDLOKQLRPYMENMPLPGI 180
Db 121 IYAPFKSLIPWNPDDIYFGGWDISNNMADAMARAKVFDIDLOKQLRPYMENMPLPGI 180
Qy 181 YPDDFIANQGSRAHVIGKTKQEOVDHIIKIDIAFKKATKYDKVYVMTANTERYSNLV 240
Db 181 YPDDFIANQGSRAHVIGKTKQEOVDHIIKIDIAFKKATKYDKVYVMTANTERYSNLV 240
Qy 241 VGLNDTMENTLAAVDRNEAEISPTLYAIACVMEVPIFNGSPONTFVPGILDLAIART 300
Db 241 VGLNDTMENTLAAVDRNEAEISPTLYAIACVMEVPIFNGSPONTFVPGILDLAIART 300
Qy 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Db 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALIVPGEHPDHVVVYKYVPYVDSKRAMDEYTSSEIFMGKSTIVLANHC 420
Db 361 NVDDMVNSNALIVPGEHPDHVVVYKYVPYVDSKRAMDEYTSSEIFMGKSTIVLANHC 420
Qy 421 EDSLAAPIILLDLVLAELSTRIEFKANEKGKFSHFVPAATILSYLTAKAPLVPGTPV 480
Db 421 EDSLAAPIILLDLVLAELSTRIEFKANEKGKFSHFVPAATILSYLTAKAPLVPGTPV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 3
T50021
inositol-3-phosphate synthase (EC 5.5.1.4) T31P16.160 [similarity] - Arabidopsis thalian

N:Alternate names: protein T31P16.160
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T50021
R:Beyan, M.; Zimmermann, W.; Grennaisen, A.; Wandut, R.; Kalicki, J.; Wohldmann, P.; Smi
submitted to the Protein Sequence Database, May 2000
A:Reference number: 225027
A:Accession: T50021
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-510 <BEV>
A:Cross-references: UNIPROT:Q9LX12; EMBL:AL356332; GSPDB:GN00063; ATSP:T31P16.160
A:Experimental source: cultivar Columbia; BAC clone T31P16
C:Genetics:
A:Gene: ATSP:T31P16.160
A:Map position: 5
A:Insertions: 63/2; 86/2; 131/3; 214/2; 290/1; 328/3; 387/3; 450/3; 471/3
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 90.0%; Score 2369; DB 2; Length 510;
Best Local Similarity 87.5%; Pred. No. 3e-152;
Matches 446; Conservative 42; Mismatches 22; Indels 0; Gaps 0;

Qy 1 MEIENFKVCEPNVKKYTEIETSIQSVNYETTELVHENRNGTYQWIVPKSKYKEFKTNHVP 60
Db 1 MFIESFKVESPNVKKYTEIENISVVDYETTELVHENRNGAFQWTVKPKYKEFKTDTRVP 60
Qy 61 KLGVWLVGWGNNGSTLTGVIANREGISMAWKIKIQOANVFGLTQASIRVGSFOGEE 120
Db 61 KLGVWLVGWGNNGSTLTGVIANREGISMAWKIKQOANVFGLTQASSIRVGSNGEE 120
Qy 121 IYAPFKSLIPWNPDDIYFGGWDISNNMADAMARAKVFDIDLOKQLRPYMESMLPLPGI 180
Db 121 IYAPFKSLIPWNPDDIYFGGWDISNNMADAMARAKVFDIDLOKQMRPMEHWPLPGI 180
Qy 121 IYAPFKSLIPWNPDDIYFGGWDISNNMADAMARAKVFDIDLOKQMRPMEHWPLPGI 180
Db 121 IYAPFKSLIPWNPDDIYFGGWDISNNMADAMARAKVFDIDLOKQMRPMEHWPLPGI 180
Qy 181 YPDDFIANQGSRAHVIGKTKQEOVDHIIKIDIAFKKATKYDKVYVMTANTERYSNLV 240
Db 181 YPDDFIANQGSRAHVIGKTKQEOVDHIIKIDIAFKKATKYDKVYVMTANTERYSNLV 240
Qy 241 VGLNDTMENTLAAVDRNEAEISPTLYAIACVMEVPIFNGSPONTFVPGILDLAIART 300
Db 241 VGLNDTMENTLAAVDRNEAEISPTLYAIACVMEVPIFNGSPONTFVPGILDLAIART 300
Qy 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Db 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALIVPGEHPDHVVVYKYVPYVDSKRAMDEYTSSEIFMGKSTIVLANHC 420
Db 361 NVDDMVNSNALIVPGEHPDHVVVYKYVPYVDSKRAMDEYTSSEIFMGKSTIVLANHC 420
Qy 421 EDSLAAPIILLDLVLAELSTRIEFKANEKGKFSHFVPAATILSYLTAKAPLVPGTPV 480
Db 421 EDSLAAPIILLDLVLAELSTRIEFKANEKGKFSHFVPAATILSYLTAKAPLVPGTPV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 4
T08436
inositol-3-phosphate synthase (EC 5.5.1.4) [similarity] - rape
C:Species: Brassica napus (rape)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T08436
R:Huseain, A.; Bourgeois, J.; Polvi, S.; Tsang, E.; Keller, W.A.; Georges, F.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z16418
A:Accession: T08436
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

A:Residues: 1-509 <HUS>
A:Cross-references: UNIPROT:Q96348; EMBL:U66307; NID:g1513227; PID:g1513228
C:Function:
A:Description: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase

Query Match 89.8%; Score 2363; DB 2; Length 509;
Best Local Similarity 88.0%; Pred. No. 7.5e-152;
Matches 448; Conservative 33; Mismatches 28; Indels 0; Gaps 0;

```

QY 2 FIENFKVCEPNVKTETETESQSVNYETTELVEHNRNGTYQWIVKPKSVKTEFTNIHVPK 61
DB 1 FIESFVESPVKYTNENHSYVDYETTEVHEHNGAYQWIVKPKSVKTEFTNIDRVK 60
QY 62 LGVNLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGEI 121
DB 61 LGVNLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFNGEEM 120
QY 122 YAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAVFDIDLOKQLRPYMESMLPLPGIY 181
DB 121 YAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAVFDIDLOKQLRPYMENIVPLPGIY 180
QY 182 DDPEFLAANQEBRANNVTKGTQEQVOQIIKDIKAFKATKVDKVVLTMTANTERYSNLVY 241
DB 181 DDPEFLAANQSRANNVTKGTKEQVDQIIKDMREFEKKNVDKVVLTMTANTERYSNLVY 240
QY 242 GLNDTMENTLAAVDNRNAEISPTLYAIACVMEVNPPIGSPONTFPGGLDLAIARNLT 301
DB 241 GLNDTMENTLMSVDRDEISPTLYAIACVLEGIPINGSPONTFPGGLDLAIKNNVL 300
QY 302 ICGDDPFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNNLSAPQTFRSKEISK 361
DB 301 ICGDDPFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNNLSAPQTFRSKEISK 360
QY 362 VVDDMNNSNALIYEPGEHPDHVVYIKVPPYVGSKRAMDEYTSIEIMGSKSTVLNHTCE 421
DB 361 VVDDMNVASNGILFEPGEHPDHVVYIKVPPYVADSKRAMDEYTSIEIMGSKNTLVNHTCE 420
QY 422 DSLLAAPILDLVLAELSTRIEPKAENEGKFSFHPVATILSYLTAKPLVPPEPTVNA 481
DB 421 DSLLAAPILDLVLAELSTRIOFKSEKGFSPHPVATKSYLTAKPLVPPEPTVNA 480
QY 482 LSKORAMLENIMRACVGLAPENNMLEYK 510
DB 481 LSKORAMLENIMRACVGLAPENNMLEYK 509

```

RESULT 5

T05017
Inositol-3-phosphate synthase (EC 5.5.1.4) T19P19.190 [similarity] - Arabidopsis thaliana
N:Alternate names: protein T19P19.190
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #ext_change 09-Jul-2004
C:Accession: T05017
R:Byvan, M.; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Hoheisel, J.; Mewes, H.W.;
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15394
A:Accession: T05017
A:Molecule type: DNA
A:Residues: 1-511 <BRV>
A:Cross-references: UNIPROT:P42801; EMBL:A1022605
A:Experimental source: cultivar Columbia; BAC clone T19P19
C:Genetics:
A:Map position: 4
A:Introns: 64/2; 87/2; 132/3; 215/2; 291/1; 329/3; 388/3; 451/3
A>Note: T19P19.190
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 89.4%; Score 2352.5; DB 2; Length 511;
Best Local Similarity 87.7%; Pred. No. 3.9e-151;
Matches 448; Conservative 36; Mismatches 26; Indels 1; Gaps 1;

```

QY 1 MEIENFKVCEPNVKTETETESQSVNYETTELVEHNR-NGTYQWIVKPKSVKTEFTNIHVP 59
DB 1 MEIESFKVESPVKYTNENHSYVDYETTEVHEKTNSTYQWIVKPKSVKTEFTNIDRVK 60
QY 60 PKLGWNLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGE 119
DB 61 PKLGWNLVGLGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFNGE 120
QY 120 EIVAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAVFDIDLOKQLRPYMESMLPLPG 179
DB 121 EIVAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAVFDIDLOKQLRPYMENIVPLPG 180
QY 180 IYDPEFLAANQEBRANNVTKGTQEQVOQIIKDIKAFKATKVDKVVLTMTANTERYSNLV 239
DB 181 IYDPEFLAANQSRANNVTKGTKEQVDQIIKDMREFEKKNVDKVVLTMTANTERYSNLV 240
QY 240 VVGLNDTMENTLAAVDNRNAEISPTLYAIACVMEVNPPIGSPONTFPGGLDLAIARN 299
DB 241 VVGLNDTMENTLMSVDRDEISPTLYAIACVLEGIPINGSPONTFPGGLDLAIARN 300
QY 300 TLIIGDDPFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNNLSAPQTFRSKEISK 359
DB 301 TLIIGDDPFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNNLSAPQTFRSKEISK 360
QY 360 SNVDDMNNSNALIYEPGEHPDHVVYIKVPPYVGSKRAMDEYTSIEIMGSKSTVLNHT 419
DB 361 SNVDDMNVASNGILFEPGEHPDHVVYIKVPPYVADSKRAMDEYTSIEIMGSKNTLVNHT 420
QY 420 CEDSLLAAPILDLVLAELSTRIEPKAENEGKFSFHPVATILSYLTAKPLVPPEPTPV 479
DB 421 CEDSLLAAPILDLVLAELSTRIOFKSEKGFSPHPVATILSYLTAKPLVPPEPTPV 480
QY 480 NALSKORAMLENIMRACVGLAPENNMLEYK 510
DB 481 NALSKORAMLENIMRACVGLAPENNMLEYK 511

```

RESULT 6

S60302
Inositol-3-phosphate synthase (EC 5.5.1.4) - Spirodeja polytricha
C:Species: Spirodeja polytricha
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #ext_change 09-Jul-2004
C:Accession: S60302
R:Smart, C.C.; Fleming, A.J.
Plant J. 4, 279-293, 1993
A:Title: A plant gene with homology to D-myo-inositol-3-phosphate synthase is rapidly an
A:Reference number: S60302; MUID:94035182; PMID:8220483
A:Accession: S60302
A:Molecule type: mRNA
A:Residues: 1-510 <SMA>
A:Cross-references: UNIPROT:P42803; EMBL:Z11693; NID:G396209; PID:CAA7751.1; PID:G55864
C:Genetics:
A:Gene: turl
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase

Query Match 89.2%; Score 2349; DB 2; Length 510;
Best Local Similarity 87.3%; Pred. No. 6.6e-151;
Matches 445; Conservative 40; Mismatches 25; Indels 0; Gaps 0;

```

QY 1 MEIENFKVCEPNVKTETETESQSVNYETTELVEHNRNGTYQWIVKPKSVKTEFTNIHVP 60
DB 1 MEIESFKVESPVKYTNENHSYVDYETTEVHEHNGAYQWIVKPKSVKTEFTNIDRVK 60
QY 61 KLGWNLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGE 120
DB 61 KLGWNLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFNGE 120
QY 121 YAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAVFDIDLOKQLRPYMESMLPLPGI 180
DB 121 YAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAVFDIDLOKQLRPYMESMLPLPGI 180

```



```
|||||
Db 421 CEDSLAARPIIDLVLLAEISTRIGKSGEGKFSFHPVATILSLTKAPLVPPTGPV 480
Qy 480 NALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 NALSKORAMLENIMRACVGLAPENNMILEYK 511

RESULT 9
T04399
inositol-3-phosphate synthase (EC 5.5.1.4) - barley
C/Species: Hordeum vulgare (barley)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #ext_change 09-Jul-2004
C/Accession: T04399
R/Larson, S.R.; Raboy, V.
Submitted to the EMBL Data Library, March 1998
A/Description: Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.
A/Reference number: Z14366
A/Accession: T04399
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-510 <LAK>
A/Cross-references: UNIPROT:O65195; EMBL:AF056325; NID:g3152730; PIDN:AC17133.1; PID:g3
A/Experimental source: cv. Harrington
C/Genetics:
A:Gene: INO1
A:Map position: 4
C/Function:
A/Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-inc
A/Pathway: inositol biosynthesis
A/Note: first step
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase; NAD

Query Match 88.4%; Score 2328; DB 2; Length 510;
Best Local Similarity 87.6%; Pred. No. 1.7e-149;
Matches 447; Conservative 26; Mismatches 37; Indels 0; Gaps 0;

Qy 1 MFIENPKVCEPVNKTETETIQSVNYYETTELHNRNGTYOMIVKRSKYKEPKTNIHP 60
Db 1 MFIENPKVCEPVNKTETETIQSVNYYETTELHNRNGTYOMIVKRSKYKEPKTNIHP 60
Qy 61 KLGVMVLVGNGNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGE 120
Db 61 KLGVMVLVGNGNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGE 120
Qy 121 IYAPFKSLIPVNPDPDIVFGWDISNMNLADAMARAVFDILOKOLRPMESMLPLPGI 180
Db 121 IYAPFKSLIPVNPDPDIVFGWDISNMNLADAMARAVFDILOKOLRPMESMLPLPGI 180
Qy 181 YPDPFLAANQERRANNVIGTKOEVOQIIKDIKAFKATKYDKVYVLTANTERSNLY 240
Db 181 YPDPFLAANQERRANNVIGTKOEVOQIIKDIKAFKATKYDKVYVLTANTERSNLY 240
Qy 241 VGLNPTMENTLAAVDRNEAEISPSLYAIACVMEVNPFGNSPONTFVGLIDLAIARNT 300
Db 241 VGLNPTMENTLAAVDRNEAEISPSLYAIACVMEVNPFGNSPONTFVGLIDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTISIVSYNHLGNNDGMNLASAPQTRSKEISKS 360
Db 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTISIVSYNHLGNNDGMNLASAPQTRSKEISKS 360
Qy 361 NVVDDMVNSNALYERBEHPDHVVVIKYVPYVGDSCRADDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALYERBEHPDHVVVIKYVPYVGDSCRADDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAEISTRIFEKAENEGKFSFHPVATILSLTKAPLVPPTGPV 480
Db 421 EDSLAAPIIIDLVLLAEISTRIFEKAENEGKFSFHPVATILSLTKAPLVPPTGPV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

```
|||||
RESULT 10
S52648
inositol-3-phosphate synthase (EC 5.5.1.4) - Citrus paradisi
C/Species: Citrus paradisi
C/Date: 19-Mar-1997 #sequence_revision 09-May-1997 #ext_change 09-Jul-2004
C/Accession: S52648
R/Holland, D.
Submitted to the EMBL Data Library, April 1994
A/Reference number: S52648
A/Accession: S52648
A/Molecule type: DNA
A/Residues: 1-507 <HOL>
A/Cross-references: UNIPROT:P42802; GB:Z32632; NID:g602564; PIDN:CAA83565.1; PID:g602565
C/Genetics:
A:Gene: INO1
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase

Query Match 87.9%; Score 2314.5; DB 2; Length 507;
Best Local Similarity 88.0%; Pred. No. 1.4e-148;
Matches 449; Conservative 30; Mismatches 28; Indels 3; Gaps 2;

Qy 1 MFIENPKVCEPVNKTETETIQSVNYYETTELHNRNGTYOMIVKRSKYKEPKTNIHP 60
Db 1 MFIENPKVCEPVNKTETETIQSVNYYETTELHNRNGTYOMIVKRSKYKEPKTNIHP 60
Qy 61 KLGVMVLVGNGNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGE 120
Db 61 KLGVMVLVGNGNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGE 120
Qy 121 IYAPFKSLIPVNPDPDIVFGWDISNMNLADAMARAVFDILOKOLRPMESMLPLPGI 180
Db 121 IYAPFKSLIPVNPDPDIVFGWDISNMNLADAMARAVFDILOKOLRPMESMLPLPGI 180
Qy 181 YPDPFLAANQERRANNVIGTKOEVOQIIKDIKAFKATKYDKVYVLTANTERSNLY 240
Db 181 YPDPFLAANQERRANNVIGTKOEVOQIIKDIKAFKATKYDKVYVLTANTERSNLY 240
Qy 241 VGLNPTMENTLAAVDRNEAEISPSLYAIACVMEVNPFGNSPONTFVGLIDLAIARNT 300
Db 241 VGLNPTMENTLAAVDRNEAEISPSLYAIACVMEVNPFGNSPONTFVGLIDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTISIVSYNHLGNNDGMNLASAPQTRSKEISKS 360
Db 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTISIVSYNHLGNNDGMNLASAPQTRSKEISKS 360
Qy 361 NVVDDMVNSNALYERBEHPDHVVVIKYVPYVGDSCRADDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALYERBEHPDHVVVIKYVPYVGDSCRADDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAEISTRIFEKAENEGKFSFHPVATILSLTKAPLVPPTGPV 480
Db 421 EDSLAAPIIIDLVLLAEISTRIFEKAENEGKFSFHPVATILSLTKAPLVPPTGPV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 478 ALSKORAMLENIMRACVGLAPENNMILEYK 507

RESULT 11
T18569
inositol-3-phosphate synthase (EC 5.5.1.4) - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jul-2004
C/Accession: T18569; T20002
R/Ainscough, R.
Submitted to the EMBL Data Library, December 1998
A/Reference number: T18569
A/Accession: T18569
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
```


Best Local Similarity 27.7%; Pred. No. 1.8e-12;
Matches 110; Conservative 62; Mismatches 157; Indels 68; Gaps 11;

```
OY      61  KLGVNLVGMGNGSTLUGVIANREGISWATKDIQQANVFSLTQASAIRVGSFQGE 120
Db      2  RTGILVGRSGSVATTVTYGAALRAGLVPAT-----GCVSALEA-----FDG-- 44
OY      121 IYAPFKSLPMVNPDDIVFGWMDISNMNL--ADAMARAKVPDIDLQKQLRPYMESMLP- 176
Db      45  -----VPLRGFDELVFGHDVGTGLVKRAEQLABAGV-----VPR 80
OY      177 -LPGIYDPDFIAANOBERANNVIKGTKQ-----EQVOQIIKDIAFKKATKVDKXY 226
Db      81  GLPGVLTAELEDAADAERIP--APGTBEGGDGTPDQASAAAGIIVADLTGFRERLGLDRYV 138
OY      227 VLTANTERYSNLVVGLNDTMENLLAADRNEAFISPTLYAIACMVENVPTINGSPQ-N 285
Db      139 VNVVSTQPPA-VPHPAHSAALREBALRGERPLPVSSLYAYALRAACAFVDFPSTG 197
OY      286 TFVPGILDLAIARNTLIGDDPFKSGQTKMKSVLVDFLVGAGIKPTISVSYNHLGNNDGMN 345
Db      198 ARLPALDELAREQGLPYAGSDGKTGETLVKSVLAMPARALRVRSWSGTNLGGDGAT 257
OY      346 LSAPQTFRSKEISKSNVVDMDVNSNALIYEPGEHPDHVVYIKYVPYVGDSCRAMDEYTS 405
Db      258 IADPERVVSKNASKGIVLE-----AELGHAVEGGVHHHVDPDLEGWKTAMDHVTFE 308
OY      406 IFMGKSTVLHNTCEDSLAAPITLIDLVLAEISTR 442
Db      309 GFLGAKMTLQFTWQGDSSLAAPLVLDLARFMALAH 345
```

Search completed: June 7, 2005, 16:47:40
Job time : 28.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:46:53 ; Search time 81.5 Seconds
(without alignments) 2254.032 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631

Sequence: 1 MFINFKVESPNVXYTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599520 seqs, 360203123 residues

Total number of hits satisfying chosen parameters: 1599520

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2631	100.0	510	US-10-025-003-12	Sequence 12, Appl
2	2631	100.0	510	US-10-718-952-12	Sequence 11, Appl
3	2624	99.7	510	US-10-025-003-14	Sequence 14, Appl
4	2624	99.7	510	US-10-025-003-16	Sequence 16, Appl
5	2624	99.7	510	US-10-718-952-14	Sequence 14, Appl
6	2624	99.7	510	US-10-718-952-16	Sequence 16, Appl
7	2593	98.6	510	US-10-025-003-2	Sequence 2, Appl
8	2593	98.6	510	US-10-025-003-10	Sequence 10, Appl
9	2593	98.6	510	US-10-718-952-2	Sequence 2, Appl
10	2593	98.6	510	US-10-718-952-10	Sequence 10, Appl
11	2588	98.4	510	US-10-025-003-6	Sequence 6, Appl
12	2588	98.4	510	US-10-718-952-6	Sequence 6, Appl

13	2582.5	98.2	511	US-10-424-599-213009	Sequence 213009, A
14	2525	96.0	510	US-10-424-599-154864	Sequence 154864, A
15	2431	92.4	505	US-10-442-017-15	Sequence 15, Appl
16	2366	89.9	510	US-10-425-115-231852	Sequence 231852, A
17	2366	89.9	516	US-10-425-114-625678	Sequence 625678, A
18	2365	89.9	510	US-10-767-701-462278	Sequence 462278, A
19	2364	89.9	510	US-10-739-930-7635	Sequence 7635, Ap
20	2361	89.7	510	US-10-425-115-206814	Sequence 206814, A
21	2355	89.5	542	US-10-425-114-58674	Sequence 58674, A
22	2352	89.4	510	US-10-437-963-173252	Sequence 173252, A
23	2346	89.2	510	US-09-921-232-11	Sequence 11, Appl
24	2346	89.2	510	US-09-921-330-11	Sequence 11, Appl
25	2346	89.2	510	US-09-921-339-11	Sequence 11, Appl
26	2337	88.8	510	US-09-727-628-2	Sequence 2, Appl
27	2337	88.8	510	US-10-425-115-231850	Sequence 231850, A
28	2337	88.8	510	US-10-425-115-231857	Sequence 231857, A
29	2337	88.8	512	US-10-425-114-39621	Sequence 39621, A
30	2337	88.8	512	US-10-425-114-53231	Sequence 53231, A
31	2337	88.8	512	US-10-425-114-66216	Sequence 66216, A
32	2336	88.8	510	US-10-425-115-231853	Sequence 231853, A
33	2336	88.8	512	US-10-425-114-66916	Sequence 46916, A
34	2221.5	84.4	509	US-10-437-963-187768	Sequence 187768, A
35	2203	83.7	510	US-10-425-115-268216	Sequence 268216, A
36	2203	83.7	521	US-10-425-114-62888	Sequence 62888, A
37	2203	83.7	522	US-10-425-114-62888	Sequence 62888, A
38	2203	83.7	529	US-10-425-114-58895	Sequence 58895, A
39	2133	81.1	458	US-10-425-115-231843	Sequence 231843, A
40	2100	79.8	450	US-10-425-115-231845	Sequence 231845, A
41	2032	77.2	431	US-10-424-599-154863	Sequence 154863, A
42	1616	61.4	558	US-09-876-889-348	Sequence 348, App
43	1612.5	61.3	534	US-10-128-714-3515	Sequence 3515, Ap
44	1608.5	61.1	534	US-10-128-714-8515	Sequence 8515, Ap
45	1558	59.2	332	US-10-425-114-42702	Sequence 42702, A

ALIGNMENTS

RESULT 1

US-10-025-003-12

Sequence 12, Application US/10025003

Publication No. US20030074685A1

GENERAL INFORMATION:

APPLICANT: Hitz, William

APPLICANT: Sebastian, Scott

APPLICANT: Grace, John

TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE

TITLE OF INVENTION: SACHARIDES AND PHYTIC ACID

FILE REFERENCE: BB-1077-C

CURRENT APPLICATION NUMBER: US/10/025, 003

CURRENT FILING DATE: 2002-05-07

PRIOR APPLICATION NUMBER: 08/835,751

PRIOR FILING DATE: APRIL 8, 1997

PRIOR APPLICATION NUMBER: PCT/US98/06822

PRIOR FILING DATE: APRIL 7, 1998

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Microsoft Office 97

SEQ ID NO 12

LENGTH: 510

TYPE: PRT

ORGANISM: Glycine max

US-10-025-003-12

Query Match 100.0%; Score 2631; DB 14; Length 510;

Best Local Similarity 100.0%; Pred. No. 9,8e-218;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MFINFKVESPNVXYTETETIOSVYNYETTELAVHNNRGTQWIVKPSVNYQFTKTHVP 60

1 MFINFKVESPNVXYTETETIOSVYNYETTELAVHNNRGTQWIVKPSVNYQFTKTHVP 60

61 KLGWVLWGNGNNSLTLCGVIANREDISWATKXKIQQANVFSLQASAIRVSGFOEE 120

```
Db 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVSGFQEE 120
121 IYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
121 IYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
Qy 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Qy 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIIDLAIANT 300
241 VGLNDTMENLLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIIDLAIANT 300
Db 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIIDLAIANT 300
Qy 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTISVSYNHLGNNDGNMLSAPOFRSKEISKS 360
301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTISVSYNHLGNNDGNMLSAPOFRSKEISKS 360
Db 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTISVSYNHLGNNDGNMLSAPOFRSKEISKS 360
Qy 361 NVVDMDMNSNAIIVPEBHPDHVVVIXVPYVGDSKRAMDEYTSSEIFMGKNTIVLHNTC 420
361 NVVDMDMNSNAIIVPEBHPDHVVVIXVPYVGDSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Db 361 NVVDMDMNSNAIIVPEBHPDHVVVIXVPYVGDSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Qy 421 EDSLLAAPTIIIDLVLALSTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTAVN 480
421 EDSLLAAPTIIIDLVLALSTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTAVN 480
Db 421 EDSLLAAPTIIIDLVLALSTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTAVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
```

RESULT 2

```
US-10-718-952-12
; Sequence 12, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-12
```

```
Query Match 100.0%; Score 2631; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 9.8e-218;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MFENFKVESPNNKYTETETIOSVYNYETTELVHENRNGTYOMIYKPSVNYQFKTNTHVP 60
1 MFENFKVESPNNKYTETETIOSVYNYETTELVHENRNGTYOMIYKPSVNYQFKTNTHVP 60
Db 1 MFENFKVESPNNKYTETETIOSVYNYETTELVHENRNGTYOMIYKPSVNYQFKTNTHVP 60
Qy 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVSGFQEE 120
61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVSGFQEE 120
Db 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVSGFQEE 120
Qy 121 IYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
121 IYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
Db 121 IYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
```

```
Qy 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Qy 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIIDLAIANT 300
241 VGLNDTMENLLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIIDLAIANT 300
Db 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIIDLAIANT 300
Qy 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTISVSYNHLGNNDGNMLSAPOFRSKEISKS 360
301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTISVSYNHLGNNDGNMLSAPOFRSKEISKS 360
Db 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTISVSYNHLGNNDGNMLSAPOFRSKEISKS 360
Qy 361 NVVDMDMNSNAIIVPEBHPDHVVVIXVPYVGDSKRAMDEYTSSEIFMGKNTIVLHNTC 420
361 NVVDMDMNSNAIIVPEBHPDHVVVIXVPYVGDSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Db 361 NVVDMDMNSNAIIVPEBHPDHVVVIXVPYVGDSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Qy 421 EDSLLAAPTIIIDLVLALSTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTAVN 480
421 EDSLLAAPTIIIDLVLALSTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTAVN 480
Db 421 EDSLLAAPTIIIDLVLALSTRIOFKAENEGKFSFHPVATILSYLTKAPLVPGTAVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
```

RESULT 3

```
US-10-025-003-14
; Sequence 14, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-14
```

```
Query Match 99.7%; Score 2624; DB 14; Length 510;
Best Local Similarity 99.8%; Pred. No. 3.9e-217;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MFENFKVESPNNKYTETETIOSVYNYETTELVHENRNGTYOMIYKPSVNYQFKTNTHVP 60
1 MFENFKVESPNNKYTETETIOSVYNYETTELVHENRNGTYOMIYKPSVNYQFKTNTHVP 60
Db 1 MFENFKVESPNNKYTETETIOSVYNYETTELVHENRNGTYOMIYKPSVNYQFKTNTHVP 60
Qy 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVSGFQEE 120
61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVSGFQEE 120
Db 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVSGFQEE 120
Qy 121 IYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
121 IYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
Db 121 IYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESVPLPGI 180
Qy 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVLMTANTERYSNLV 240
Qy 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIIDLAIANT 300
241 VGLNDTMENLLAAVDRNEAISPSTLYAIACWENVFPINGSPONTFVPGIIDLAIANT 300
```


Db 241 VGLNDTMENTLLAAVDRNEAEISPSSTLYAIACMENVPPFINSPOQTFPVGILDLAIANT 300
Qy 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTISIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
Db 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTISIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Qy 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
Db 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
Qy 481 ALSKORAMLENIMRAACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRAACVGLAPENNMILEYK 510

RESULT 4

US-10-025-003-16
; Sequence 16, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-16

Query Match 99.7%; Score 2624; DB 14; Length 510;

Best Local Similarity 99.8%; Pred. No. 3.9e-217;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MFIEHFVESPNVKTETEIOGVNYETTELVEHNRNGTYQMIYVPKSVNTQFKTNTHP 60
Db 1 MFIEHFVESPNVKTETEIOGVNYETTELVEHNRNGTYQMIYVPKSVNTQFKTNTHP 60
Qy 61 KLGVWLVMGNGNSTLTGVIANREGISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Db 61 KLGVWLVMGNGNSTLTGVIANREGISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPWNPPDIFYFGGMDISNMNLADAMARAKYFDIDLOKOLRPYMESWVPLFGI 180
Db 121 IYAPFKSLPWNPPDIFYFGGMDISNMNLADAMARAKYFDIDLOKOLRPYMESWVPLFGI 180
Qy 121 IYAPFKSLPWNPPDIFYFGGMDISNMNLADAMARAKYFDIDLOKOLRPYMESWVPLFGI 180
Db 121 IYAPFKSLPWNPPDIFYFGGMDISNMNLADAMARAKYFDIDLOKOLRPYMESWVPLFGI 180
Qy 181 YDPDFIAANOEBRANNVIKGTKEOVQOIKIDIKAFKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANOEBRANNVIKGTKEOVQOIKIDIKAFKATKVDKVVLMTANTERYSNLV 240
Qy 181 YDPDFIAANOEBRANNVIKGTKEOVQOIKIDIKAFKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANOEBRANNVIKGTKEOVQOIKIDIKAFKATKVDKVVLMTANTERYSNLV 240
Qy 241 VGLNDTMENTLLAAVDRNEAEISPSSTLYAIACMENVPPFINSPOQTFPVGILDLAIANT 300
Db 241 VGLNDTMENTLLAAVDRNEAEISPSSTLYAIACMENVPPFINSPOQTFPVGILDLAIANT 300
Qy 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTISIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
Db 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTISIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420

Qy 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Qy 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
Db 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
Qy 481 ALSKORAMLENIMRAACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRAACVGLAPENNMILEYK 510

RESULT 5

US-10-718-952-14
; Sequence 14, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-14

Query Match 99.7%; Score 2624; DB 16; Length 510;

Best Local Similarity 99.8%; Pred. No. 3.9e-217;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MFIEHFVESPNVKTETEIOGVNYETTELVEHNRNGTYQMIYVPKSVNTQFKTNTHP 60
Db 1 MFIEHFVESPNVKTETEIOGVNYETTELVEHNRNGTYQMIYVPKSVNTQFKTNTHP 60
Qy 61 KLGVWLVMGNGNSTLTGVIANREGISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Db 61 KLGVWLVMGNGNSTLTGVIANREGISWATKDKIOQANFGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPWNPPDIFYFGGMDISNMNLADAMARAKYFDIDLOKOLRPYMESWVPLFGI 180
Db 121 IYAPFKSLPWNPPDIFYFGGMDISNMNLADAMARAKYFDIDLOKOLRPYMESWVPLFGI 180
Qy 121 IYAPFKSLPWNPPDIFYFGGMDISNMNLADAMARAKYFDIDLOKOLRPYMESWVPLFGI 180
Db 121 IYAPFKSLPWNPPDIFYFGGMDISNMNLADAMARAKYFDIDLOKOLRPYMESWVPLFGI 180
Qy 181 YDPDFIAANOEBRANNVIKGTKEOVQOIKIDIKAFKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANOEBRANNVIKGTKEOVQOIKIDIKAFKATKVDKVVLMTANTERYSNLV 240
Qy 181 YDPDFIAANOEBRANNVIKGTKEOVQOIKIDIKAFKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANOEBRANNVIKGTKEOVQOIKIDIKAFKATKVDKVVLMTANTERYSNLV 240
Qy 241 VGLNDTMENTLLAAVDRNEAEISPSSTLYAIACMENVPPFINSPOQTFPVGILDLAIANT 300
Db 241 VGLNDTMENTLLAAVDRNEAEISPSSTLYAIACMENVPPFINSPOQTFPVGILDLAIANT 300
Qy 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTISIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
Db 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTISIVSYNHLGNDGNMLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
Qy 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480
Db 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPPGTPPVN 480

Db 421 EDSLAABIIDVLVLAELSTRIOFKAENEGKFSFHVATILSYLTAKAPLVPGTPTVVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 6
US-10-718-952-16

Sequence 16, Application US/10718952
Publication No. US20040128713A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT FILING DATE: 2003-11-21
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO: 16
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-718-952-16

Query Match 99.7%; Score 2624; DB 16; Length 510;
Best Local Similarity 99.8%; Pred. No. 3.9e-217;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MFIEFKVESPVVKYETETIETIOSVYNYETTELVEHNRNGTQYMIYKPSVNYQFKNTHTVP 60
Db 1 MFIEFKVESPVVKYETETIETIOSVYNYETTELVEHNRNGTQYMIYKPSVNYQFKNTHTVP 60
Qy 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIOQANYFGSLTQASAIRVGSFOGEE 120
Db 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIOQANYFGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKVFDIDLOKQRPYMESVPLPGI 180
Db 121 IYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKVFDIDLOKQRPYMESVPLPGI 180
Qy 181 YDPDFIAANOBERANNVIKGTQEQVOOIKDIAFKKATKVDKVVVMTANTERYSNLY 240
Db 181 YDPDFIAANOBERANNVIKGTQEQVOOIKDIAFKKATKVDKVVVMTANTERYSNLY 240
Qy 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACVMENVPFINGSPONTFVPGILDLAIANT 300
Db 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACVMENVPFINGSPONTFVPGILDLAIANT 300
Qy 301 LIIGDDFSGQTKKMSVVDLVGAGIKPTISIVSYNHLGNDGNLSPQTFRSKEISKS 360
Db 301 LIIGDDFSGQTKKMSVVDLVGAGIKPTISIVSYNHLGNDGNLSPQTFRSKEISKS 360
Qy 361 NVVDVMNSNALIYEPGEHPDHVVIVKYVPYVGDSCRAMDEYSEIFMGKNTIYLAHNTC 420
Db 361 NVVDVMNSNALIYEPGEHPDHVVIVKYVPYVGDSCRAMDEYSEIFMGKNTIYLAHNTC 420
Qy 421 EDSLAABIIDVLVLAELSTRIOFKAENEGKFSFHVATILSYLTAKAPLVPGTPTVVN 480
Db 421 EDSLAABIIDVLVLAELSTRIOFKAENEGKFSFHVATILSYLTAKAPLVPGTPTVVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 7

US-10-025-003-2

Sequence 2, Application US/10025003
Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT FILING DATE: 2002-05-07
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO: 2
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-025-003-2

Query Match 98.6%; Score 2593; DB 14; Length 510;
Best Local Similarity 98.4%; Pred. No. 1.8e-214;

Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MFIEFKVESPVVKYETETIETIOSVYNYETTELVEHNRNGTQYMIYKPSVNYQFKNTHTVP 60
Db 1 MFIEFKVESPVVKYETETIETIOSVYNYETTELVEHNRNGTQYMIYKPSVNYQFKNTHTVP 60
Qy 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIOQANYFGSLTQASAIRVGSFOGEE 120
Db 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIOQANYFGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKVFDIDLOKQRPYMESVPLPGI 180
Db 121 IYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKVFDIDLOKQRPYMESVPLPGI 180
Qy 181 YDPDFIAANOBERANNVIKGTQEQVOOIKDIAFKKATKVDKVVVMTANTERYSNLY 240
Db 181 YDPDFIAANOBERANNVIKGTQEQVOOIKDIAFKKATKVDKVVVMTANTERYSNLY 240
Qy 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACVMENVPFINGSPONTFVPGILDLAIANT 300
Db 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACVMENVPFINGSPONTFVPGILDLAIANT 300
Qy 301 LIIGDDFSGQTKKMSVVDLVGAGIKPTISIVSYNHLGNDGNLSPQTFRSKEISKS 360
Db 301 LIIGDDFSGQTKKMSVVDLVGAGIKPTISIVSYNHLGNDGNLSPQTFRSKEISKS 360
Qy 361 NVVDVMNSNALIYEPGEHPDHVVIVKYVPYVGDSCRAMDEYSEIFMGKNTIYLAHNTC 420
Db 361 NVVDVMNSNALIYEPGEHPDHVVIVKYVPYVGDSCRAMDEYSEIFMGKNTIYLAHNTC 420
Qy 421 EDSLAABIIDVLVLAELSTRIOFKAENEGKFSFHVATILSYLTAKAPLVPGTPTVVN 480
Db 421 EDSLAABIIDVLVLAELSTRIOFKAENEGKFSFHVATILSYLTAKAPLVPGTPTVVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 8

US-10-025-003-10

Sequence 10, Application US/10025003
Publication No. US20030074685A1

```
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-10

Query Match          98.6%; Score 2593; DB 14; Length 510;
Best Local Similarity 98.4%; Pred. No. 1.8e-214;
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MFIEFKYSPNVKKTETETIOSVNYETTELVEHNRNGTYQWIVPKSVNTQFKTNHVP 60
    |||||
DB 1 MFIEFKYECNPVKKTETETIOSVNYETTELVEHNRNGTYQWIVPKSVKTEFKTNIHVP 60
    |||||

QY 61 KLGWVLVGMGNGNSTLTGVIYANREDISMAWKDIOQANFGSLTQASAIRVGSFQGBE 120
    |||||
DB 61 KLGWVLVGMGNGNSTLTGVIYANREGISMAWKDIOQANFGSLTQASAIRVGSFQGBE 120
    |||||

QY 121 IYAPFKSLPMVNPDPDIFGGMWDSNNMLADAMARAKVFDIDLOKQLRPYMESMPLPGI 180
    |||||
DB 121 IYAPFKSLPMVNPDPDIFGGMWDSNNMLADAMARAKVFDIDLOKQLRPYMESMPLPGI 180
    |||||

QY 121 IYAPFKSLPMVNPDPDIFGGMWDSNNMLADAMARAKVFDIDLOKQLRPYMESMPLPGI 180
    |||||
DB 121 IYAPFKSLPMVNPDPDIFGGMWDSNNMLADAMARAKVFDIDLOKQLRPYMESMPLPGI 180
    |||||

QY 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIAKFKATKVDKVVVMTANTERYSNLV 240
    |||||
DB 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIAKFKATKVDKVVVMTANTERYSNLV 240
    |||||

QY 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIAKFKATKVDKVVVMTANTERYSNLV 240
    |||||
DB 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIAKFKATKVDKVVVMTANTERYSNLV 240
    |||||

QY 241 VGLNDTMENLLAAVDRNBAEISPSLTVAIACMVNPVPINSPOPTFVPGILDIAIANT 300
    |||||
DB 241 VGLNDTMENLLAAVDRNBAEISPSLTVAIACMVNPVPINSPOPTFVPGILDIAIANT 300
    |||||

QY 241 VGLNDTMENLLAAVDRNBAEISPSLTVAIACMVNPVPINSPOPTFVPGILDIAIANT 300
    |||||
DB 241 VGLNDTMENLLAAVDRNBAEISPSLTVAIACMVNPVPINSPOPTFVPGILDIAIANT 300
    |||||

QY 301 LIGGDDFKSGOTKMSVLVDLVGAGIKPTISIVSYNHLGNDGMNLSAPQTFRSKEISKS 360
    |||||
DB 301 LIGGDDFKSGOTKMSVLVDLVGAGIKPTISIVSYNHLGNDGMNLSAPQTFRSKEISKS 360
    |||||

QY 301 LIGGDDFKSGOTKMSVLVDLVGAGIKPTISIVSYNHLGNDGMNLSAPQTFRSKEISKS 360
    |||||
DB 301 LIGGDDFKSGOTKMSVLVDLVGAGIKPTISIVSYNHLGNDGMNLSAPQTFRSKEISKS 360
    |||||

QY 361 NVVDDMVNSNAIIVPEGHPDHVVVVKVVPYVGSKRAMDEYTSSEIFMGKNTIVLANTC 420
    |||||
DB 361 NVVDDMVNSNAIIVPEGHPDHVVVVKVVPYVGSKRAMDEYTSSEIFMGKNTIVLANTC 420
    |||||

QY 361 NVVDDMVNSNAIIVPEGHPDHVVVVKVVPYVGSKRAMDEYTSSEIFMGKNTIVLANTC 420
    |||||
DB 361 NVVDDMVNSNAIIVPEGHPDHVVVVKVVPYVGSKRAMDEYTSSEIFMGKNTIVLANTC 420
    |||||

QY 421 EDSLAAPIIIDLVLLAELSTRIOFKANBEKPHSFHVATILSYLTARAPLVPGTPVNV 480
    |||||
DB 421 EDSLAAPIIIDLVLLAELSTRIOFKANBEKPHSFHVATILSYLTARAPLVPGTPVNV 480
    |||||

QY 421 EDSLAAPIIIDLVLLAELSTRIOFKANBEKPHSFHVATILSYLTARAPLVPGTPVNV 480
    |||||
DB 421 EDSLAAPIIIDLVLLAELSTRIOFKANBEKPHSFHVATILSYLTARAPLVPGTPVNV 480
    |||||

QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
    |||||
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
    |||||

RESULT 9
US-10-718-952-2
; Sequence 2, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
```

```
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-2

Query Match          98.6%; Score 2593; DB 16; Length 510;
Best Local Similarity 98.4%; Pred. No. 1.8e-214;
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MFIEFKYSPNVKKTETETIOSVNYETTELVEHNRNGTYQWIVPKSVNTQFKTNHVP 60
    |||||
DB 1 MFIEFKYECNPVKKTETETIOSVNYETTELVEHNRNGTYQWIVPKSVKTEFKTNIHVP 60
    |||||

QY 61 KLGWVLVGMGNGNSTLTGVIYANREDISMAWKDIOQANFGSLTQASAIRVGSFQGBE 120
    |||||
DB 61 KLGWVLVGMGNGNSTLTGVIYANREGISMAWKDIOQANFGSLTQASAIRVGSFQGBE 120
    |||||

QY 121 IYAPFKSLPMVNPDPDIFGGMWDSNNMLADAMARAKVFDIDLOKQLRPYMESMPLPGI 180
    |||||
DB 121 IYAPFKSLPMVNPDPDIFGGMWDSNNMLADAMARAKVFDIDLOKQLRPYMESMPLPGI 180
    |||||

QY 121 IYAPFKSLPMVNPDPDIFGGMWDSNNMLADAMARAKVFDIDLOKQLRPYMESMPLPGI 180
    |||||
DB 121 IYAPFKSLPMVNPDPDIFGGMWDSNNMLADAMARAKVFDIDLOKQLRPYMESMPLPGI 180
    |||||

QY 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIAKFKATKVDKVVVMTANTERYSNLV 240
    |||||
DB 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIAKFKATKVDKVVVMTANTERYSNLV 240
    |||||

QY 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIAKFKATKVDKVVVMTANTERYSNLV 240
    |||||
DB 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIAKFKATKVDKVVVMTANTERYSNLV 240
    |||||

QY 241 VGLNDTMENLLAAVDRNBAEISPSLTVAIACMVNPVPINSPOPTFVPGILDIAIANT 300
    |||||
DB 241 VGLNDTMENLLAAVDRNBAEISPSLTVAIACMVNPVPINSPOPTFVPGILDIAIANT 300
    |||||

QY 241 VGLNDTMENLLAAVDRNBAEISPSLTVAIACMVNPVPINSPOPTFVPGILDIAIANT 300
    |||||
DB 241 VGLNDTMENLLAAVDRNBAEISPSLTVAIACMVNPVPINSPOPTFVPGILDIAIANT 300
    |||||

QY 301 LIGGDDFKSGOTKMSVLVDLVGAGIKPTISIVSYNHLGNDGMNLSAPQTFRSKEISKS 360
    |||||
DB 301 LIGGDDFKSGOTKMSVLVDLVGAGIKPTISIVSYNHLGNDGMNLSAPQTFRSKEISKS 360
    |||||

QY 301 LIGGDDFKSGOTKMSVLVDLVGAGIKPTISIVSYNHLGNDGMNLSAPQTFRSKEISKS 360
    |||||
DB 301 LIGGDDFKSGOTKMSVLVDLVGAGIKPTISIVSYNHLGNDGMNLSAPQTFRSKEISKS 360
    |||||

QY 361 NVVDDMVNSNAIIVPEGHPDHVVVVKVVPYVGSKRAMDEYTSSEIFMGKNTIVLANTC 420
    |||||
DB 361 NVVDDMVNSNAIIVPEGHPDHVVVVKVVPYVGSKRAMDEYTSSEIFMGKNTIVLANTC 420
    |||||

QY 361 NVVDDMVNSNAIIVPEGHPDHVVVVKVVPYVGSKRAMDEYTSSEIFMGKNTIVLANTC 420
    |||||
DB 361 NVVDDMVNSNAIIVPEGHPDHVVVVKVVPYVGSKRAMDEYTSSEIFMGKNTIVLANTC 420
    |||||

QY 421 EDSLAAPIIIDLVLLAELSTRIOFKANBEKPHSFHVATILSYLTARAPLVPGTPVNV 480
    |||||
DB 421 EDSLAAPIIIDLVLLAELSTRIOFKANBEKPHSFHVATILSYLTARAPLVPGTPVNV 480
    |||||

QY 421 EDSLAAPIIIDLVLLAELSTRIOFKANBEKPHSFHVATILSYLTARAPLVPGTPVNV 480
    |||||
DB 421 EDSLAAPIIIDLVLLAELSTRIOFKANBEKPHSFHVATILSYLTARAPLVPGTPVNV 480
    |||||

QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
    |||||
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
    |||||

RESULT 10
US-10-718-952-10
; Sequence 10, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
```

```

: PRIOR APPLICATION NUMBER: PCT/US98/06832
: PRIOR FILING DATE: APRIL 7, 1998
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: Microsoft Office 97
: SEQ ID NO: 10
: LENGTH: 510
: TYPE: PRT
: ORGANISM: Glycine max
: US-10-718-952-10

```

Query Match	Score	DB	Length
98.6%	2593	16	510

Qy	1	MFLENFVESPNTKYETETELQS VYNNVETTEL VHENNGTYQVI VPKSVNYOFKNTIHP	60
Dp	1	MFLENFVESPNTKYETETELQS VYNNVETTEL VHENNGTYQVI VPKSVNYOFKNTIHP	60
Qy	61	KLGVMVLVGMGNGNSLTLTGVTIANREDISWATYDKQIQOANYGSLTQMSAIRVGSFOGEE	12
Dp	61	KLGVMVLVGMGNGNSLTLTGVTIANREGISWATYDKIQOANYGSLTQMSAIRVGSFOGEE	12
Qy	121	IYAPKSLIPMNPDDIVGGMQDISMNNIADAMARAQVDDIDLOQLRPYESSMPLRGI	18
Dp	121	IYAPKSLIPMNPDDIVGGMQDISMNNIADAMARAQVDDIDLOQLRPYESSMPLRGI	18
Qy	181	YDPDFIAANGEBRANNVIKGTQKQOYQOIIKDIKAFKATKYDVYVMTANTERYSNLV	24
Dp	181	YDPDFIAANGEBRANNVIKGTQKQOYQOIIKDIKAFKATKYDVYVMTANTERYSNLV	24
Qy	241	VGNDDMTENMLAANDRENEAISPSTLYALACWENPFINSGPQTFPBGILDLAIANT	30
Dp	241	VGNDDMTENMLAANDRENEAISPSTLYALACWENPFINSGPQTFPBGILDLAIANT	30
Qy	301	LIGGDDFKSGQTRKMSLYVDFLVGAGIKPETSIVSYNHLGNNGMMLSAPQTFRSKEISKS	36
Dp	301	LIGGDDFKSGQTRKMSLYVDFLVGAGIKPETSIVSYNHLGNNGMMLSAPQTFRSKEISKS	36
Qy	361	NVVDDMWNNAIIYEGEHPDHVVVKKVPPYGDSSKRADEYTSSEIFMGKRTIYLANHC	42
Dp	361	NVVDDMWNNAIIYEGEHPDHVVVKKVPPYGDSSKRADEYTSSEIFMGKRTIYLANHC	42
Qy	421	EDSLIAPRILDLVLAELSTRIOFPAENEGKHSFHPATILSYLTKAPLVPQTPPVVN	48
Dp	421	EDSLIAPRILDLVLAELSTRIEFPAENEGKHSFHPATILSYLTKAPLVPQTPPVVN	48
Qy	481	ALSKORAMLENIMRACVGLAPENNMILEXK 510	
Dp	481	ALSKORAMLENIMRACVGLAPENNMILEXK 510	

```

11      RESULT 11
12      US-10-025-003-6
13      ; Sequence 6, Application US/10025003
14      ; Publication No. US20030074685A1
15      ; GENERAL INFORMATION:
16      ; APPLICANT: Hitz, William
17      ; APPLICANT: Sebastian, Scott
18      ; APPLICANT: Grace, John
19      ; APPLICANT: Streitz, Leon
20      ; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
21      ; FILE REFERENCE: BB-1077-C
22      ; CURRENT APPLICATION NUMBER: US/10/025,003
23      ; CURRENT FILING DATE: 2002-05-07
24      ; PRIOR APPLICATION NUMBER: 08/835,751
25      ; PRIOR FILING DATE: APRIL 8, 1997
26      ; PRIOR APPLICATION NUMBER: PCT/US98/06822
27      ; PRIOR FILING DATE: APRIL 7, 1998
28      ; NUMBER OF SEQ ID NOS: 16
29      ; SOFTWARE: Microsoft Office 97
30      ; SEQ ID NO 6
31      ; LENGTH: 510

```

```

; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-6

```

Query Match	98.4%	Score 2588	DB 14	Length 510
Best Local Similarity	98.2%	Pred. No. 5e-214		
Matches 501	Conservative 4	Mismatches 5	Indels 0	Gaps 0

Qy	1	MFLFNKVECPNVKTEHETIEIOSVYVYETTELVEHERNNTQYIMKPKSPVYQOFKTNHVP	60
Qy	1	MFLFNKVECPNVKTEHETIEIOSVYVYETTELVEHERNNTQYIMKPKSPVYQOFKTNHVP	60
Qy	1	MFLFNKVECPNVKTEHETIEIOSVYVYETTELVEHERNNTQYIMKPKSPVYQOFKTNHVP	60
Qy	61	KLGCVMLVGGNGNGSTLTGCVJANFEDISMATKDKIQOANTFGSLTOASAIRVGSFGGEE	120
Qy	61	KLGCVMLVGGNGNGSTLTGCVJANFEGISMATKDKIQOANTFGSLTOASAIRVGSFGGEE	120
Qy	121	IYAPFKSLLLPMVNPDDIYFGGWDISNNMLAAMARAKYFDIDLOKOPRPYMESVPLPGI	180
Qy	121	IYAPFKSLLLPMVNPDDIYFGGWDISNNMLAAMARAKYFDIDLOKOPRPYMESVPLPGI	180
Qy	181	YDDBFTAAQOEERANVIKGTQOEVOOIIKDKIKAFKATYDKVYVLMJANTERYSNLV	240
Qy	181	YDDBFTAAQOEERANVIKGTQOEVOOIIKDKIKAFKATYDKVYVLMJANTERYSNLV	240
Qy	241	VGGNDTMENTLLAAYVRNEAIEISPTLYAIACMEVNEPPINGSPONTFVPGJIDLAIRNT	300
Qy	241	VGGNDTMENTLLAAYVRNEAIEISPTLYAIACMEVNEPPINGSPONTFVPGJIDLAIRNT	300
Qy	301	LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNNDGNMLSAPOTKRSKEISKS	360
Qy	301	LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNNDGNMLSAPOTKRSKEISKS	360
Qy	361	NVYDDMVNSNATILYEBGEHPDHVYIKXVPVYVGSKPMDEYTSIEIFMGKNNTVLHNTC	420
Qy	361	NVYDDMVNSNATILYEBGEHPDHVYIKXVPVYVGSKPMDEYTSIEIFMGKNNTVLHNTC	420
Qy	421	EDSLLAAPITLLDVLAEILSTRIOFKAENEGKFHSFHPVATILSYLTAKPLVPBPGTEVN	480
Qy	421	EDSLLAAPITLLDVLAEILSTRIOFKAENEGKFHSFHPVATILSYLTAKPLVPBPGTEVN	480
Qy	481	ALSROBAMLENTMRACVGLAPENNMLILEYK	510
Qy	481	ALSROBAMLENTMRACVGLAPENNMLILEYK	510

```

RESULT 12
US-10-718-952-6
, Sequence 6, Application US/10718952
, Publication No. US20040128713A1
, GENERAL INFORMATION:
, APPLICANT: Hitz, William
, APPLICANT: Sebastian, Scott
, APPLICANT: Grace, Tom
, APPLICANT: Steelf, Leon
, TITLE OF INVENTION: SOYBAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
, TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
, FILE REFERENCE: BB-1077-C
, CURRENT APPLICATION NUMBER: US/10/718,952
, CURRENT FILING DATE: 2003-11-21
, PRIOR APPLICATION NUMBER: 08/835,751
, PRIOR FILING DATE: APRIL 8, 1997
, PRIOR APPLICATION NUMBER: PCT/US98/06822
, PRIOR FILING DATE: APRIL 7, 1998
, NUMBER OF SEQ ID NOS: 16
, SOFTWARE: Microsoft Office 97
, SEQ ID NO 6
, LENGTH: 510
, TYPE: PRT
, ORGANISM: Glycine max
US-10-718-952-6

```

Query Match	98.4%;	Score 2588;	DB 16;	Length 510
Best Local Similarity	98.2%;	Pred. No. 5e-214;		

Matches 501; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 MFIEFKVESPNVKTETETESVNYETTELVEHNRNGTYQWIVPKSVNTQFKTNTHP 60
Db 1 MFIEFKVCECPVKTETETESVNYETTELVEHNRNGTYQWIVPKSVKYEFKTNIHVP 60

QY 61 KLGVWLVMGNGNSTLTGGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGE 120
Db 61 KLGVWLVMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGE 120

QY 121 IYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESWVPLPGI 180
Db 121 IYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESWVPLPGI 180

QY 181 YDPDFIAANOEBERANNVIKGTKEOVQOIIKDIKAFKATKVDKVVVMTANTERYSNL 240
Db 181 YDPDFIAANOEBERANNVIKGTKEOVQOIIKDIKAFKATKVDKVVVMTANTERYSNL 240

QY 241 VGLNDTMENLLAAVDRNEAISPSTLYAACWENVPFINGSPONTFVPGILDIAIARN 300
Db 241 VGLNDTMENLLAAVDRNEAISPSTLYAACWENVPFINGSPONTFVPGILDIAIARN 300

QY 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTISVYNHLGNDGMNLSAPQTFRSKEISK 360
Db 301 LIGGDDFKSGOTKMSVLDVFLVAGIKPTISVYNHLGNDGMNLSAPQTFRSKEISK 360

QY 361 NVVDDMVNSNAIIEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLANT 420
Db 361 NVVDDMVNSNAIIEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLANT 420

QY 421 EDSLILAAPILIDVLLAELSTRIOFKAENEGKTHSFHPVATILSYLTRKAPLVPPTPVV 480
Db 421 EDSLILAAPILIDVLLAELSTRIEFKAENEGKTHSFHPVATILSYLTRKAPLVPPTPVV 480

QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 13
US-10-424-599-213009
; Sequence 213009, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 213009
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(511)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1.pep
US-10-424-599-213009

Query Match 98.2%; Score 2582.5; DB 15; Length 511;
Best Local Similarity 98.2%; Pred. No. 1.5e-213;
Matches 502; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 1 MFIEFKVESPNVKTETETESVNYETTELVEHNRNGTYQWIVPKSVNTQFKTNTHP 60
Db 1 MFIEFKVCECPVKTETETESVNYETTELVEHNRNGTYQWIVPKSVKYEFKTNIHVP 60

QY 61 KLGVWLVMGNGNSTLTGGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGE 119
Db 61 KLGVWLVMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGE 120

QY 120 EYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESWVPLPGI 179
Db 121 EYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESWVPLPGI 180

QY 180 IYDPDFIAANOEBERANNVIKGTKEOVQOIIKDIKAFKATKVDKVVVMTANTERYSNL 239
Db 181 IYDPDFIAANOEBERANNVIKGTKEOVQOIIKDIKAFKATKVDKVVVMTANTERYSNL 240

QY 240 VGLNDTMENLLAAVDRNEAISPSTLYAACWENVPFINGSPONTFVPGILDIAIARN 299
Db 241 VGLNDTMENLLAAVDRNEAISPSTLYAACWENVPFINGSPONTFVPGILDIAIARN 300

QY 300 TLIGGDDFKSGOTKMSVLDVFLVAGIKPTISVYNHLGNDGMNLSAPQTFRSKEISK 359
Db 301 TLIGGDDFKSGOTKMSVLDVFLVAGIKPTISVYNHLGNDGMNLSAPQTFRSKEISK 360

QY 360 SNVDDMVNSNAIIEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLANT 419
Db 361 SNVDDMVNSNAIIEPGEHPDHVVVIKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLANT 420

QY 420 CEDSLILAAPILIDVLLAELSTRIOFKAENEGKTHSFHPVATILSYLTRKAPLVPPTPVV 479
Db 421 CEDSLILAAPILIDVLLAELSTRIEFKAENEGKTHSFHPVATILSYLTRKAPLVPPTPVV 480

QY 480 NALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 NALSKORAMLENIMRACVGLAPENNMILEYK 511

RESULT 14
US-10-424-599-154864
; Sequence 154864, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154864
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110863C.1.pep
US-10-424-599-154864

Query Match 96.0%; Score 2525; DB 15; Length 510;
Best Local Similarity 94.7%; Pred. No. 1.3e-208;
Matches 483; Conservative 18; Mismatches 9; Indels 0; Gaps 0;

QY 1 MFIEFKVESPNVKTETETESVNYETTELVEHNRNGTYQWIVPKSVNTQFKTNTHP 60
Db 1 MFIEFKVCECPVKTETETESVNYETTELVEHNRNGTYQWIVPKSVKYEFKTNIHVP 60

QY 61 KLGVWLVMGNGNSTLTGGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGE 120
Db 61 KLGVWLVMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGE 120

QY 121 IYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESWVPLPGI 180
Db 121 IYAPFKSLPVMNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKOLRPYMESWVPLPGI 180

Qy	181	YDPEFIANOEERANNYIKGTROEOVOOIIDIRAFKATQVQVYVMTANTERSYLV	240
Db	181	YDPEFIANQCDKRAITWIRKGTKEQVOOIIDIEFKESTQVQVYVMTANTERSYLV	240
Qy	241	VGINDMENTLIAAVDRNEAEISPSSTLYAIACWENVPINGSPONTVPGLIDIALARNT	300
Db	241	VGINDTMBENFASIDRNEAEISPSSTLYAIACWENVPINGSPONTFPGIDIALIKRNS	300
Qy	301	LIGGDDFSGGOTKKKSUYVDFLYGAGIKPTSIIVSYNHGNNDDGNLSAPQTFRSKEISKS	360
Db	301	LIGGDDFKSGOTKKKSUYVDFLYGAGIKPTSIIVSYNHGNNDDGNLSAPQTFRSKEISKS	360
Qy	361	NVDDDMNSNAIYEPGEHPDHVVYIKVYPVYGDSKRAMDEYTSBIFMGKNNTVLNATC	420
Db	361	NVDDDMNSNAIYEPGEHPDHVVYIKVYPVYGDSKRAMDEYSSBIFMGKNNTVLNATC	420
Qy	421	EDSLIAPPIIDVLVLAELSTRIQFKAENEGSKFHSFHVATILSYLTAPLVPPEGTPVYN	480
Db	421	EDSLIAPPIIDVLVLAELSTRIEFKAENEGSKFHSFHVATILSYLTAPLVPPESTPVYN	480
Qy	481	ALSKORAMLENTIRACVGLAPENNMTILEYK 510	
Db	481	ALSKORAMLENTIRACVGLAPENNMTILEYK 510	

```

RESULT 15
US-10-442-017-15
; Sequence 15, Application US/10442017
; Publication No. US20030215950A1
; GENERAL INFORMATION:
; APPLICANT: Battelle Memorial Institute
; APPLICANT: Laesure, Linda L.
; APPLICANT: Dai, Ziyu
; TITLE OF INVENTION: Isolated Polynucleotides and Methods of Promoting a Morphology In
; TITLE OF INVENTION: a Fungus
; FILE REFERENCE: BA4-195
; CURRENT APPLICATION NUMBER: US/10/442, 017
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: 60/382,132
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 505
; TYPE: PRF
; ORGANISM: Sesamum indicum
; US-10-442-017-15

```

Query Match	92.4%;	Score 2431;	DB 15;	Length 505;
Best Local Similarity	92.3%;	Pred. No. 1.7e-200;		
Matches 466;	Conservative 19;	Mismatches 20;	Indels 0;	Gaps 0

Qy	6	KKASPNVKTETELQSYVNYETTELVEHNGNGYQMTVVKRSYVNOQKTNTHPKJGVM	65
Db	1	FKASSPNVKTETEGEHSYVNYETTELVEHSGNGYQMTVVKRYKYERKJTNTHPKJGVM	60
Qy	66	LVGNGNGNSTLTVGCVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFQGEETIYAPF	125
Db	61	LVGNGNGNSTLTVGCVIANREGISWATKDKIQOANYFGSLTQASSIRVGSFNGEETIYAPF	120
Qy	126	KSLLPMVNPDDIVFVGQMDISNNMNLADAMARAKAFVFDIDLOKOLRPFMESEMVPLPGIYDPDF	185
Db	121	KSLLPMVNPDDIVFVGQMDISNNMNLADAMGRAKVFDIDLOKOLRPFMEHNVPLPGIYDPDF	180
Qy	186	IAANOEBRANNVIKGTQEOVQOIIKDIKAFKEATKVDKVVVLVMTANTERYSNLVYGIND	245
Db	181	IAANOGBRANNVIKGTKEQVOQIHKDMRDREQNKVDKVVVLVMTANTERSNVVVGIND	240
Qy	246	TMENILAAVDNEABEISPTLYAIAACVMEVNPFIKSGPONTFVPLDILAIARNTLIGGD	305
Db	241	TABSLMASVEBNEABEISPTLYAIAACVPEVNPFIKSGPONTFVPLDILAIQNRSLIGGD	300
Qy	306	DFKSGQTRKMSVLVDFLVGAGIKPETSIVSYNHLGNNDGMSLAPOTFRSKEISKSNNVD	365

Db	301	DFASGTRKAKSLVDFPLVGAGIKPISIVSYNHLGNDGMLSLAQPTRSKSEISKSANVDD	360
Qy	366	MNSNNAIIYEPGEHPDHVVIKYVPYVDSKRAMDEYTSSEIFMGKNTIVLHNTCEBSLL	425
Db	361	MAANSNGIYEPGEHPDHVVIKYVPYVDSGRAMDEYTSSEIFMGKNTIVLHNTCEBSLL	420
Qy	426	AAPIIIDVLVLAEISTRIOFKAENGEKHSFHPVATILSYLTKAPLVPFGTPVNNALSKQ	485
Db	421	AAPIIIDVLVLAEISTRIOFKAENGEKHSFHPVATILSYLTKAPLVPFGTPVNNALSKQ	480
Qy	486	RAMLENIMRACVGLAPENNMILEYK	510
Db	481	RAMLENIMRACVGLAPENNMILEYK	505

Search completed: June 7, 2005, 17:03:08
Job time : 82.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 7, 2005, 17:11:57 ; Search time 5315 Seconds
(without alignments)
4649.513 Million cell updates/sec

Title: US-10-718-952-2
Perfect score: 2632
Sequence: 1 MFIEHPKVCPCNVKKTETETI.....NIMBACVGLAPENMILEYK 510

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+.p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10718952/runat_06062005.173400.12846/app.query.fasta_1.1294
-DB=GenEmbl1 -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdt -LIT=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=act -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10718952.OCEN_1_17010@runat_06062005.173400.12846 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl1:
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Query Match	Length DB ID	Description
1	2632	100.0 1739 8	AY038802 Glycine m
2	2632	100.0 1782 6	BD075266 Soybean p
3	2637	99.8 1533 6	BD075269 Soybean p
4	2619	99.5 1791 8	AF293970 Glycine m

5	2477	94.1	1959	8	AB009881	AB009881 Nicotiana
6	2472	93.9	1931	8	AB032073	AB032073 Nicotiana
7	2472	93.9	1950	6	E27176	E27176 Novel INPS
8	2469	93.8	1959	8	BT013505	BT013505 Lycopersi
9	2453	93.2	1845	8	AF284065	AF284065 Sesamum 1
10	2439	92.7	1954	8	BT013759	BT013759 Lycopersi
11	2392	90.9	1538	8	AF120146	AF120146 Trilicium
12	2392	90.9	1538	8	AF120147	AF120147 Trilicium
13	2392	90.9	1538	8	AF120148	AF120148 Trilicium
14	2382	90.9	1966	8	AF433879	AF433879 Suaeda ma
15	2381	90.5	2053	8	MCU32511	U32511 Mesembryant
16	2378	90.3	1533	6	C0805008	C0805008 Sequence
17	2378	90.3	1533	6	AX0506743	AX0506743 Sequence
18	2378	90.3	1533	8	AY143904	AY143904 Arabidops
19	2378	90.3	1863	8	AY054202	AY054202 Arabidops
20	2378	90.3	1864	8	AY053415	AY053415 Arabidops
21	2374.5	90.2	1845	8	AY028259	AY028259 Avicennia
22	2369	90.0	1564	8	AY096554	AY096554 Arabidops
23	2369	90.0	1863	8	AY065415	AY065415 Arabidops
24	2368	90.0	1781	6	AX054630	AX054630 Sequence
25	2368	90.0	1781	8	BN066307	BN066307 Brasica na
26	2356	89.5	1533	8	AY323824	AY323824 Xerophyta
27	2355	89.5	1890	8	ATU30250	ATU30250 Arabidops
28	2352.5	89.4	1567	8	BT001931	BT001931 Arabidops
29	2352.5	89.4	1837	8	AY085931	AY085931 Arabidops
30	2352.5	89.4	1870	8	AF372954	AF372954 Arabidops
31	2349	89.2	1842	8	SPW1P8SYM	Z11693 S.polyrrhiz
32	2347	89.2	1914	8	AK103501	AK103501 Oryza sat
33	2347	89.2	1915	8	AK058750	AK058750 Oryza sat
34	2344	89.1	1535	6	BD073472	BD073472 Regulated
35	2344	89.1	1665	6	BD073470	BD073470 Regulated
36	2344	89.1	1665	8	AF056326	AF056326 Zea mays
37	2340	88.9	1868	8	AB012107	AB012107 Oryza sat
38	2338.5	88.8	1921	8	ATU04876	U04876 Arabidops
39	2337.5	88.8	1978	8	CPTN0048	Z32682 C.paradis
40	2336	88.8	1931	6	AR137808	AR137808 Sequence
41	2336	88.8	1931	6	AR170167	AR170167 Sequence
42	2335	88.7	1959	6	AB151207	AB151207 Sequence
43	2333	88.6	1936	8	AB059557	AB059557 Avena sat
44	2331.5	88.6	1871	8	PV038920	U38920 Phaseolus v
45	2338	88.4	2152	8	AF056325	AF056325 Hordeum v

ALIGNMENTS

RESULT 1	AY038802	1739 bp	mRNA	linear	PLN 16-JUN-2001
LOCUS	Glycine max myo-inositol-1-phosphate synthase mRNA, complete cds.				
DEFINITION	AY038802				
ACCESSION	AY038802.1	GI:14764465			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Glycine max (soybean)				
REFERENCE	1 (bases 1 to 1739)				
AUTHORS	Hitz,W.D., Carlson,T.J., Kerr,P. and Sebastian,S.				
TITLE	Biochemical and Molecular Characterization of a Mutation that Confers a Decreased Rafinosaecharide and Phytic Acid Phenotype on Soybean Seeds				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1739)				
AUTHORS	Carlson,T.J. and Hitz,W.D.				
TITLE	Direct Substitution				
JOURNAL	Submitted (06-JUN-2001) Crop Genetics, DuPont Co., P.O.Box 80402, Wilmington, DE 19880-0402, USA				
FEATURES	Location/Qualifiers				
source	1..1739				
	/organism="Glycine max"				
	/mol_type="mRNA"				

/cultivar="Mye"
/db_xref="taxon:3847"
/dev_stage="developing embryo"
11..1543
/EC_number="5.5.1.4"
/function="cyclize glucose 6-phosphate to
L-myo-inositol-1-phosphate"
/note="mIPS-1A"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAK72098.1"
/db_xref="GI:1476446"

CDS

/translation="MPIENFKYECNPKVKTETEIQSVYNYETTELHENRNTYOMIV
KPKSVYERKTNIIHVPLKGVMLVGMGNNSTLTGVLANREGISWATKDKIOANVF
GSUTQASAIRVCSFOGEYERPAFKSLIPVNPDDIVFGMDISMNILADAMARVFD
IDLQKOLRPMESMLPPLGCTYDPDFLAAQBERANNVKGTRQEOVOQIITKOITAPKE
ARKVDKVVLTWANTERYSNLVGLNDTMENLAAVDREBAEISPTLVALACWENV
PFINSGPQTFVPGILDIAIARNTLIGDDPKSGQTKMSKLVDFLVGAGIKPTISVS
YHNLGNNDMDLAPQTFRSKEISKSNVVDVNSNALIYEPGEHDHVYVKKYVYV
GDSKRAMDEYTSRIFMGKSTIVLHNTGDSILAAPILDLVLAELSTRIFKANE
CKFHSFHPVATLLSLITKAPLVYPCGIPVYNALSKORAMLENIMRACVGLAPENNMLE
YK"

ORIGIN

Alignment Scores:

Pred. No.: 3,74e-201 Length: 1739
Score: 2632.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-718-952-2 (1-510) x AY038802 (1-1739)

QY 1 MetPheIleGluAsnPhelYsValGluCySProAsnValIlyTyThrGluThrgluIle 20
DB 11 ATGTCATCGAGATTTTAAAGTTGAGTGTCCTAAGTGAAGTACACCGAGACTGAGATT 70
QY 21 GlnSerValTyRAsnTyrgIuThrThrgIuEuValHiGluAsnArgAsnGlyThrTyR 40
DB 71 CAGTCGGTGTACACATACGAAACCACTGTTACGAGAAACGGAATGACACTTAT 130
QY 41 GlnTrpIleValIlySerProIlySerValIlyTyrgIuPhelYsThraSnIleHisValPro 60
DB 131 CAGTGGATGTGCAAAACCAATCTGTCAAATACGAATTTAAACCAACATCCATGTTCT 190
QY 61 LysLeuGlyValMetLeuValGlyTrpGlyGlyAsnAsnGlySerThrLeuThrgIyGly 80
DB 191 AATTAGGGGTATATGCTTGTGGGTGGGATGGAACCAACGGCTCAACCGGTGGT 250
QY 81 ValIleAlaAsnArgGluGlyIleSerTrpAlaThrIlyAsnIlyIleGlnGlnAlaAsn 100
DB 251 GTTATTGCTAACGAGGAGGCAATTCATGGGCTAACAGAACAAAGATTCAACAGCCAA 310
QY 101 TyRPhelGlySerLeuThrgIuAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 311 TACTTTGGCTCCCTCAACGCTCACTATCCGAGTTGGGCTTCCAGGAGAGGAA 370
QY 121 IleTyRAlaProPhelYsSerLeuLeuProMetValAsnProAsnAspIleValPheGly 140
DB 371 ACTATAGCCCATTCAGAGCGCTTCCAAATGTGTAAACCTGACAGACATTTGTTGGG 430
QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIlyValPheAsp 160
DB 431 GGAATGGGAATATACGAAATGAACCTGGCTGATGCATGCGCCAGGCAAAAGGTTTGAC 490
QY 161 IleAspLeuGlnIlySerGlnLeuArgProTyRMetGlySerMetLeuProLeuProGlyIle 180
DB 491 ATCGATTTGCAAGAGAGTTGAGGCTTTCATGGAATCCATGTTCCATCCCGGAATC 550
QY 181 TyRAspProAspPheIleAlaAsnGlnGluGlyIlyArgAlaAsnAsnValIleIlyGly 200
DB 551 TATGACCCGGAATTTCTGCTGCCAACCAAGAGGAGCGTGCCAAACATCATCAAGGCG 610

QY 201 ThrIlySGlnGluGlnValGlnGlnIleIleIlyAspIleIleYsAlaPheIlySGluAlaThr 220
DB 611 ACMAAGCAAGAGCAAGTTCACAAATCATCAAGACATCAAGGCTTTAAGAAAGCCACC 670
QY 221 IlyValIAspIlyValIValIleuTrpThraIAsnThrgIuArgTySerAsnLeuVal 240
DB 671 AAAGTGACAAAGGTGGTGTACTGTGAGCTGCCAACACAGAGAGTACAGTAAATTTGGTT 730
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaIValAspArgAsnGluIleGlu 260
DB 731 GTGGGCTTTAATGACACCATGGAATCTTGGCTGTGTGACAGAAATAGGCTGAG 790
QY 261 ILeSerProSerThrLeuTyRAlaIleAlaCyValIleGluAsnValProPheIleAsn 280
DB 791 ATTTCTCTTCCACCTTGATGCCATTCCTGTGTATGGAAGAAAGTTCCTTCAATTAT 850
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAnthr 300
DB 851 GGAAGCCCTCAGAACACTTTGTACAGAGGCTGATTGATCTTGCCATCGCGAGAACT 910
QY 301 LeuIleGlyIlyAspAspPheIlySerGlyIuThrIlyMetIlySerValIleuValAsp 320
DB 911 TTGATTTGGTGAGATGACTTCAAGATGCTCAGACCAAAATGAATCTGTGTGGTTGAT 970
QY 321 PheLeuValGlyAlaGlyIleIlySerProThrSerIleValSerTyRAsnHisLeuGlyAsn 340
DB 971 TTCCTTGTGGGGGCTGCTGATACAGCAACATCTATAGTACAGTTACACCATCTGGGAAAC 1030
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIlyGluIleSerIlySer 360
DB 1031 AATGATGGTAAATCTTGGGCTCCACAACTTCCGTTCCAGAAATCTCCCAAGAGC 1090
QY 361 AsnValIValAspAspMetValAsnSerAsnAlaIleuTyrgIuProGlyGluHisPro 380
DB 1091 AACGTTGTGAGAAATGATGTCACACAGCAATGCAATCTCTATAGCCTGGGTAAACATCCA 1150
QY 381 AspHisValIValIleIlySerTyRAlaProTyRValGlyAspSerIlyArgAlaMetAsp 400
DB 1151 GACCATGTTGTTGTTATTAAGATGTGCTTACGTAGGGAGACAGCAAGAACCATGAT 1210
QY 401 GlyTyRThrSerGluIlePheMetGlyGlyIlySerThrIleValIleuHisAnthrCyS 420
DB 1211 GAGTACACTTCAGAGATATTATTCAGGTGGAAAGACACCATGTTTGGCAACAACATGC 1270
QY 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleuLeuAlaGluLeuSer 440
DB 1271 GAGGATTCCTCTTAGCTGCTCTATATCTTGGACTTGTCTTGTGCTAGCTCAGC 1330
QY 441 ThrArgIleGluPheIlyAlaGluAsnGluGlyIlyPheHisSerPheHisProValAla 460
DB 1331 ACTAGATATGAGTTTAAAGCTAAATAGAGGAAATTCATCTCATCCACCAAGTTGCT 1390
QY 461 ThrIleLeuSerTyRLeuThrIlyAlaProLeuValProProGlyThrProValIAlaAsn 480
DB 1391 ACCATCTCAGCTTACCTCCCAAGGCTCCTGTGTTCCACCGGATACCAAGTGGGTAAT 1450
QY 481 AlaLeuSerIlyGlnArgAlaMetLeuGluAsnIleMetArgAlaCyValAlaGlyLeuAla 500
DB 1451 GCATTGTCAAAACACGCGCAATGCTGGAACCAATTAAGAGGCTTGTGTGATTTGGCC 1510
QY 501 ProGluAsnAsnMetIleuGluTyRlyS 510
DB 1511 CAGAGAAATATACATGATTTCCAGATACAG 1540

RESULT 2
BD075266 1782 bp DNA linear PART 27-AUG-2002
LOCUS
DEFINITION Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid.
ACCESSION
BD075266
BD075266.1 GI:22620869
VERSION
JP 2001519665-A/1.

SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1782)
AUTHORS Hitz, W.D. and Sebastian, S.A.
TITLE Soybean plant producing seeds with reduced levels of raffinose
JOURNAL Patent: JP 2001519665-A 1 23-OCT-2001;
E1 DU PONT DE NEMOURS AND CO
COMMENT OS Soybean line LR13
PN JP 2001519665-A/1
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN
PC C12N15/52, C12N15/82, C12N15/11, C12N9/90, A01H5/00 CC
Strandness: Double;
CC Topology: Linear;
CC Soybean plant producing seeds with reduced
levels of raffinose
CC Saccharides
CC and phytic acid
FH Key Location/Qualifiers
FT CDS 54..1586.
1..1782 Location/Qualifiers
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:
Pred. No.: 3, 85e-201 Length: 1782
Score: 2632.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-718-952-2 (1-510) x BD075266 (1-1782)

QY 1 MetPheIIeGluAsnPhelValGluCysProAsnValIleThrGluThrGluIle 20
DB 54 ATGTTTCATGAGAAATTTTAAAGTTGAGTGTCTTAATGTAGAGACCCGAGACTGAGATT 113
QY 21 GlnSerValIleValIlePheProIleValIleValIleValIleValIleValIle 40
DB 114 CAGTCGCTGACACACGAAACCGAACTTGTTCAGAGAACGAGATGCACTAT 173
QY 41 GlnTTPIIleValIlePheProIleValIleValIleValIleValIleValIle 60
DB 174 CAGTGCATGTCAAACCCAAATCTGTCAATATCGAATTTAAACCAATCCATGTTCCCT 233
QY 61 LysLeuGluValMetLeuValGluTTPGluValAsnAsnGluSerThrLeuThrGluIle 80
DB 234 AAATTAAGGGGTAATGCTTGAGGTGGGGTGGAGAAACAGGCTCAACCTCACCGGTGGT 293
QY 81 ValIleAlaAsnArgGluIleIleSerTPAlaThrIleAspIleIleGlnIleAlaAsn 100
DB 294 GTTATTGCTAACCGAGAGGCAATTCATGGGGCTACAAAGACAAAGATTCAACAGCCAAAT 353
QY 101 TyrPheGluSerLeuThrGlnAlaSerAlaIleArgValIleGlySerPheGlnIleGluIle 120
DB 354 TACCTTTGGCTCCCTCAACGAGCTCAAGCTATCCGAGTTGGGCTCTTCCAGGAGAGGAA 413
QY 121 IleTyrAlaProPheIleSerLeuLeuProMetValIleAsnProAspAspIleValIlePheGlu 140
DB 414 ATCTATGCCCATTCAGAGCTGCTTCCAAATGATTAAACCTGACGACAAATGTTGGG 473
QY 141 GlyTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValIlePheAsp 160
DB 474 GGATGGGATATACGAAACATGAACCTGGCTGATGCCATGGCCAGGCAAGGTGTTGAC 533

QY 161 IleAspLeuGlnIleValIleLeuArgProTyrMetGluSerMetLeuProLeuProGluIle 180
DB 534 ATCGATTTCAGAACAGATTGAGGCTTTCATGAGAAATCATCTTCCTCCCGGAATC 593
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluIleValIleValIleValIleValIle 200
DB 594 TATGACCCGGAATTTCTATCTCTCCCAACGAGAGGCGGTCCAAACACTCATCAAGGCG 653
QY 201 ThrIleGlnGluIleValIleGlnIleIleValIleValIleValIleValIleValIle 220
DB 654 ACAAGAGCAAGACAGATTCAACAAATCATCAAGATCATCAAGGCTTTAAGCAAGCCACC 713
QY 221 LysValIleAspLeuValIleValIleThrPheIleAsnThrGluArgTyrSerAsnLeuVal 240
DB 714 AAAGTGACAAAGGTGTTGACTGTGAGCTGCACACAGAGAGTACGATTAATTTGGTT 773
QY 241 ValIleGluLeuAsnAspThrMetGluAsnLeuLeuAlaValIleAspArgAsnGluIleGlu 260
DB 774 GTGGGCTTAAATGACACCATGAGAAATCTTGGGCTGTGACAGAAATGAGGCTGAG 833
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 834 ATTTCTCTTCCACCTTGATGCCATGTGCTGTATGAGAAATGTTCTTTCAATTAAT 893
QY 281 GlysSerProGlnAsnThrPheValIleProGluLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 894 GGAAGCCCTCAGAAACCTTTGTACAGGCTGATGATCTTGCCATCCGAGAACACT 953
QY 301 LeuIleGluValIleAspAspPheIleSerGluIleThrIleMetIleSerValIleValIleAsp 320
DB 954 TTGATTGGTGAGATGACCTTCAGAGTGTCTCAGACCAAAATGAATCTGTGTGTTGAT 1013
QY 321 PheLeuValIleValIleValIleValIleValIleValIleValIleValIleValIle 340
DB 1014 TTCCTTGGGGGCTGATATCAAGCCAACTATATATGATCAATCAACCATCTGGGAAAC 1073
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleGluIleSerIleSer 360
DB 1074 AATGATGATGATAAATCTTGAGCTCACAACTTCCTCCAGAGAAATCTCCAGAGAC 1133
QY 361 AsnValIleAspAspMetValIleAsnSerAsnAlaIleLeuTyrGluProGluIleIlePro 380
DB 1134 AACGTTGTTGATGATATGTCACAGCAATGCCATCTCTATGAGCTGTGTAACATCCA 1193
QY 381 AspHisValIleValIleValIleValIleValIleValIleValIleValIleValIle 400
DB 1194 GACCATGTTGTTTATTATTAATATGTCCTTACGTAGGGGACAGCAAGAGCCCATGAT 1253
QY 401 GluTyrThrSerGluIlePheMetGluIleValIleValIleValIleValIleValIle 420
DB 1254 GAGTACACTTCAGAAATATTCATGGGTGGAAAGACACCATTTGTTGCACAAACATGC 1313
QY 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleLeuValIleGluLeuSer 440
DB 1314 GAGGATTCCTCTTACGCTCTCTATATCTTGAGCTTGCTCTTCTGTACGCTCAGC 1373
QY 441 ThrArgIleGluPheValIleValIleValIleValIleValIleValIleValIleValIle 460
DB 1374 ACTAGAAATGAAATTAAGCTGAAATAATGAGGAAATTCACATCTCAACCCAGTTGCT 1433
QY 461 ThrIleLeuSerTyrLeuThrIleValIleValIleValIleValIleValIleValIleValIle 480
DB 1434 ACCATCTCTCAAGTACCTCAACAAAGCTCTCTGTGTTCCACCGGATACCACTGTGTAAT 1493
QY 481 AlaLeuSerIleGluArgAlaMetLeuGluAsnIleMetArgAlaCysValIleGluLeuAla 500
DB 1494 GCATTGTCAAGACAGCTGCAATGCTGAAAAACATATGAGGCTTGTGTTGATTGGCC 1553
QY 501 ProGluAsnAsnMetIleLeuGluTyrIleValIleValIleValIleValIleValIleValIle 510
DB 1554 CCAGAGAAATTAACATGATTTCTGAGTACAAAG 1583

RESULT 3

BD075269 1533 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid.
ACCESSION BD075269
VERSION BD075269.1 GI:22620872
KEYWORDS JP 2001519665-A/4.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1533)
AUTHORS Hitz,W.D. and Sebastian,S.A.
TITLE Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid
JOURNAL Patent: JP 2001519665-A 4 23-OCT-2001;
EI DU PONT DE NEMOURS AND CO
COMMENT OS Soybean line LR33
PN JP 2001519665-A/4
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ,SCOTT ANTHONY SEBASTIAN
PC C12N15/52,C12N15/82,C12N15/11,C12N9/90,A01H5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Soybean plant producing seeds with reduced
levels of raffinose
CC and phytic acid
FH Key Location/Qualifiers
FT CDS 1..1533.
Location/Qualifiers
1..1533
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:

Pred. No.:	8 05e-201	Length:	1533
Score:	2627.00	Matches:	509
Percent Similarity:	99.80%	Conservative:	0
Best Local Similarity:	99.80%	Mismatches:	1
Query Match:	99.81%	Indels:	0
DB:	6	Gaps:	0

US-10-718-952-2 (1-510) x BD075269 (1-1533)

QY 1 MetPheileGluasnPhelysValGluCySProasnValIysTyThrGluThrGluile 20
DB 1 ATGTTATCGAGAAATTTTAAGGTTGAGTGCTTAATGTGAAGTACCGAGACTGAGATT 60
QY 21 GlnSerValItyrAsnTyrglUthrThrGluLeuValHisGluAsnArgasnGlyThrTy 40
DB 61 CAGTCCGCTGACAAACGAAACCAACGAACTGTGTACAGAGAACGGAATGGCACTTAT 120
QY 41 GlnTrpIleValIysProIysSerValIysTyrgluPhelysThrAsnIleHisValPro 60
DB 121 CAGTGGATGTCAAACCAAAATCTGTCAATACGAATTTAAACCAACATCCATGTTCTT 180
QY 61 IysLeuGlyValMetLeuValGlyTrpGlyIysAsnGlySerThrLeuThrGlyGly 80
DB 181 AATTATGGGATATGCTTGAGGTTGGGGTGGGAAACACGGGCTCAACCTCACCGGTGCT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATATGCTAACGAGGCAATTTCAATGGGCTAACAGACAAAGTTCAACAGCCAAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTTGGCTCCTTACCCAGCCACGCTCAGCTATCCAGTTGGGCTCTTCCAGGAGAGGAA 360
QY 121 IleTyrrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140

DB 361 ATCTATGCCCATTCAGAGGCTGCTTCAATGGTTAACTCCCTGACGACATTTGTTGGG 420
QY 141 GYTTPAapIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGATATCAGCAACATGAACCTGGCTGATGTGACAGGCCAGGCAAGGTTGTTGAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrmecGlySerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTTGCAGAAACACATGTAGAGCTTACATGAAATCCATGCTTCCACTCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluValGlnAlaAsnValIleLysGly 200
DB 541 TATGACCCGGATTTTATTGTTGCTGCCAACAGAGGAGCGCTGCCAACAGTCATCAAGGCG 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 601 ACAAAGCAAGAGCAAGTTCAACAAATCATCAAGCATCAAGCGCTTTAAGGAAACCCACC 660
QY 221 LysValAspLysValValIleLeuTrpThrAlaAsnThrGluArgTySerAsnLeuVal 240
DB 661 AAAGTGCAAGAGTGTTGTTGTACTGTGACCTGCAACACAGAGGTACAGTATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuValAlaValAspArgAsnGluAlaGlu 260
DB 721 GTGGGCTTTAATGACACCATGAGAACTCTTGCTGCTGTGACAGAAATGAGGCTGAG 780
QY 261 IleserProserThrLeuTyrrAlaIleAlaCySValMetGluAsnValProPheIleAsn 280
DB 781 ATTCTCTTCACCTTGATATGCCATTTGCTGTGTATGAGAAATGTTCTTCTTCAATTAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAACCCCTCAAGAACATTTTGTATCAGAGGCGTGAATGATCTTGCCATCGCGAGAACACT 900
QY 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnTrpLysMetLysSerValLeuValAsp 320
DB 901 TTGATTTGGTGGAGATGACTTCAAGAGTGTGCAGCAAAATGAATCTGTGTGGTTGAT 960
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrrAsnHisLeuGlyAsn 340
DB 961 TTCCTTGGGGGGCTGGTATCAGGCCAACATCTAATAGCTATACACATCTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1021 AATGATGGTATGAATCTTCCGCTCCACAACTTCCGTTCCAAAGAAATCTCCAAAGGC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrrGluProGlyGluHisPro 380
DB 1081 AACGTTGTTGATGATATGATGTCAAACGACATGCCATCTCTATGACCTGGTGAACTCCA 1140
QY 381 AspHisValValValIleLysTrpValProTyrrValGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCATGTTGTTGTTATTAAGTATGATGCTTACGTAAGGAGACAGCAATAGACCTGAT 1200
QY 401 GlnTyrrThrSerGluIlePheMetGlyLysSerThrIleValIleuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGATATTTATGAGGTGGAAGACACACATTTGTTTGCACAAACACATGC 1260
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTCTTGGCTGCTCTATTAATCTTGGACTTGGCTCTTGGCTGAGCTCAGC 1320
QY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAGAAATCAGTTAAAGCTGAAATAGAGGAAATTCACATCTATTCACACCACTTGGCT 1380
QY 461 ThrIleLeuSerTyrrLeuThrLysAlaProLeuValProProGlyTyrrProValValAsn 480
DB 1381 ACCATCTCAGGTACCTACCAAGGCTCTCTGTTTCCACGGGGTACCAAGTGGTGAAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCySValGlyLeuAla 500

Db 1441 GCATGTCAAGACGCGTGCATGCTGAAACATATATGAGCGCTTGTTGGATTGGCC 1500

QY 501 PROGLuAsnMetIleLeuGluTyrIvs 510
|||||
Db 1501 CCAGAGATTAACATGATTCGAGTACCAAG 1530

RESULT 4
AF293970
LOCUS AF293970 1791 bp mRNA linear PLN 03-MAY-2001
DEFINITION Glycine max myo-inositol-3-phosphate synthase (MIPS) mRNA, complete cds.
ACCESSION AF293970
VERSION AF293970.1 GI:13936690
KEYWORDS Glycine max (soybean)
SOURCE Glycine max (soybean)
ORGANISM Glycine max (soybean)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 1791)
Hegeman, C.E., Good, L.L. and Grabau, E.A.
Expression of D-myo-inositol-3-phosphate synthase in soybean.
Implications for phytic acid biosynthesis
Plant Physiol. 125 (4), 1941-1948 (2001)
MEDLINE 21196082
PUBMED 11299372
REFERENCE 2 (bases 1 to 1791)
Hegeman, C.E., Good, L.L. and Grabau, E.A.
Direct Submission
Submitted (07-AUG-2000) Plant Pathology, Physiology, and Weed Science, Virginia Tech, Fralin Biotechnology Center, Blacksburg, VA 24060, USA

FEATURES
source location/Qualifiers
1. 1791
/organism="Glycine max"
/mol_type="mRNA"
/culivar="Williams 82"
/db_xref="taxon:3847"
/dev_stage="immature cotyledons"
1. 1791
/gene="MIPS"
62.. 1594
/gene="MIPS"
/note="isomerase"
/codon_start=1
/product="myo-inositol-3-phosphate synthase"
/protein_id="AAK49896.1"
/db_xref="GI:13936691"
/translation="MTENFKYECNPVKTTEBISGVNYETTELVHENRGTQWIV
KRSKYVERKTNIHVKLGVMLVGWGNGNSTLTGVIANRSGISWATKDKIQANVF
GSLTQASAIRVGSFOGEIYAPFKSLPMVNPDDIVFGMDISNMNLADAMARAKVF
IDLOKQMRYMESMLPLPGIYDPDFAAQEBRANNVIGTKOEOVOQIIXDKAKKE
ATKVKVAVLMTANTERSYNLVGLNDTMENTLLAADREARISPTIYAIACWENV
PRLNSPQRTFVPGIDLAIAARNTLGGDPKSGQTKMSYLVDPVGGIKPTSVS
VYHLNNDKMSUAPQTFRSKISKNVDDWNSVALIYEGSEHDVAVIKPYEV
GUSKRAMDEYTSIEFMGKSTVLNHTCDSLLAAPIILDVLNELSTRIFPKENE
GKFSHPHVAITLISYTKAPLVPPGTGPNALSKORAMENIMRACVGAPENNMILE
YK"

ORIGIN
Alignment Scores:
Pred. No.: 4.26e-200 Length: 1791
Score: 2619.00 Matches: 508
Percent Similarity: 99.61% Conservative: 0
Best Local Similarity: 99.61% Mismatches: 2
Query Match: 99.51% Indels: 0
DB: 8 Gaps: 0
US-10-718-952-2 (1-510) x AF293970 (1-1791)

QY 1 MetPheIleGluAsnPheIysValGluCysProAsnValIleTyrThrGluThrGluIle 20
|||||
|||||

Db 62 ATGTCATGAGAAATTTTAAGGTTGAGTGTCTATGTGAAGTACACCGAGCATGAGATT 121

QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
|||||
Db 122 CAGTCGGGTGTAACAACATCAAGAAACCAACGAACTTGTTCAAGAGAACGGAATGCGACTAT 181

QY 41 GlnTPRILEValIysPheIysSerValIysTyrGluPheIysThrAsnIleHisValPro 60
|||||
Db 182 CAGTGGATTTGTCNAACCCAAATCTGTCAAAATCGAAATTTAAACCAACATCCATGTTCT 241

QY 61 LysLeuGluValMetLeuValGlyTyrGlyAsnAsnGlySerThrLeuThrGlyGly 80
|||||
Db 242 AATTAGGGGTATAGCTTGTGGGTTGGGTTGGGAACAAGGCTCAACCTCACCGGTGGT 301

QY 81 ValIleAlaAsnArgGluGlyIleSerTPRAlaThrLysAspLysIleGlnGlnAlaAsn 100
|||||
Db 302 GTTATGCTTAACCGAGAGGCGCATTTATGGGCTTCAAAAGAACAGATTCACCAAGCCAAAT 361

QY 101 TyrPheGlySerLeuThrGlnIleAsnAlaIleArgValGlySerPheGlnGlyGlu 120
|||||
Db 362 TACTTGGCTCCCTCACAGCCTTCAGCTATCCAGTTGGGTCTTCCAGGAGAGAA 421

QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
|||||
Db 422 ATCTATGCCCATTTCAAGAGCCTGCTTCCAATGGTTAACCTGACGACATTTGTTGGG 481

QY 141 GlyTPRAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
|||||
Db 482 GGATGGGATATCAGCAACATGAACCTGCTGATGTCATGCCAGGCAAGGATTTGAC 541

QY 161 IleAspLeuGlnGlyGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyIle 180
|||||
Db 542 ATCGATTTCAAGAGAGGAGGAGGCTTACATGAATTCATGCTTCACTCCCGGAATC 601

QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluArgAlaAsnAsnValIleLysGly 200
|||||
Db 602 TATGACCGCGATTTCTATCTCTCCCAACCAAGAGAGCGGCCAACACCTCATCAAGGCG 661

QY 201 ThrLysGlnGluValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
|||||
Db 662 ACAAGACAGAACCAAGTTCACAAATCATCAAGACATCAAGCGCTTTAAGAACGACAC 721

QY 221 LysValAspLysValValIleLeuThrPheAlaAsnThrGluArgTyrSerLeuVal 240
|||||
Db 722 AAGTGAACAGAGTGTGTACTGTGACGTGCACACAGAGGTACGTAAATTTGTT 781

QY 241 ValGlyLeuAspAspThrMetGluAsnLeuLeuAlaIleValAspArgAsnGluIle 260
|||||
Db 782 GTGGGCTTATGACACCATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 841

QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
|||||
Db 842 ATTTGCTCTTCAACCTTGATGCCATGCTGTGTGTTATGAAAGAAAGTTCTCTTCAATAT 901

QY 281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
|||||
Db 902 GGAAGCCCTCAAGAACCTTTTATCACAGGCGGATGATGATCTTCCATCCCGAGAACACT 961

QY 301 LeuIleGlyValAspAspPheLysSerGlyIleThrLysMetLysSerValLeuValAsp 320
|||||
Db 962 TTGATTTGGTGAAGATTTTCAAGAGTGTCAAGACCAAAATGAATCTGTGTTGGTAT 1021

QY 321 PheLeuValGlyAlaGlyIleLysPheThrSerIleValSerTyrAsnHisLeuGlyAsn 340
|||||
Db 1022 TTCCTTGTGGGGCTGTGATCAAGCAACATCTATAGTCAGTTAACACCATCTCGGAAAC 1081

QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
|||||
Db 1082 AATGATGTATGAATCTTGGCTCCACAACTTTCCTCCAAAGAAATCTTCCAAAGAC 1141

QY 361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
|||||
Db 1142 AACGTTTGTGATGATGTGTCAACAGCATGCAATCCTCTATGAGCCTGTGTAAATCTCA 1201

QY 381 AsphHisValValIleIleYsTYrValProTYrValGlyAspSerLysArgAlaMetAsp 400
 DB 1202 GACCACTGTTGTTATTAAGATATGCTTACGTATGAGGACAGACAGAACCAATGAT 1261
 QY 401 GltYrTYrThrSerGluIlePheMetGlyGlyLysSerThrIleValIleuHisAnthrCYs 420
 DB 1262 GAGTACACTTCAAGACATATTCATGCGTGGAAAGACACCATTTGTCACCAACATGTC 1321
 QY 421 GlnAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
 DB 1322 GAGGATTCCTCTTACCTCTCTCTATTCATTCGAGATGAGTCTCTTCTTCTGAGCTCAC 1381
 QY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
 DB 1382 ACTAGATATGAGTTTAACTGAAATGAGGAAATTCACATCATTCACACCACTGCT 1441
 QY 461 ThrIleLeuSerTYrLeuThrLysAlaProLeuValProProGlyTYrThrProValValAsn 480
 DB 1442 ACCATATCTACGCTACCTCACCAAGGCTCCTGTGTTCCACCGGTCACCGGGTGTGAAT 1501
 QY 481 AlaLeuSerLysGluArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
 DB 1502 GCAATGTCAAAGCAGCTCAATGCTGAAACATATATGAGGCTTGTGTGATTGGATCGCC 1561
 QY 501 ProGluAsnAsnMetIleLeuGluTYrLys 510
 DB 1562 CCAGAGATATACATGATTTCTCGATACAG 1591
 RESULT 5
 AB009881 1959 bp mRNA linear PLN 13-JUN-2000
 LOCUS Nicotiana tabacum mRNA for myo-inositol 1-phosphate synthase,
 DEFINITION complete cds.
 ACCESSION AB009881.1 GI:8096265
 VERSION myo-inositol 1-phosphate synthase.
 KEYWORDS Nicotiana tabacum (common tobacco)
 SOURCE Nicotiana tabacum
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asteridae; lamiales; Solanales; Solanaceae; Nicotiana.
 REFERENCE 1 (sites)
 AUTHORS Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.
 TITLE Screening of wound-responsive genes identifies an immediate-early
 expressed gene encoding a highly charged protein in mechanically
 wounded tobacco plants
 JOURNAL Plant Cell Physiol. 41 (6), 684-691 (2000)
 MEDLINE 20399434
 PUBMED 10945337
 REFERENCE 2 (bases 1 to 1959)
 AUTHORS Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.
 TITLE Direct Submission
 JOURNAL Submitted (18-DEC-1997) Kojiro Hara, Nara Institute of Science and
 Technology, Plant Molecular Breeding, Takayama 891e-5, Ikoma, Nara
 630-0101, Japan (E-mail:k-hara@nbs.aisr-nara.ac.jp,
 Tel:+81-743-72-5653 (ex.5653), Fax:+81-743-72-5659)
 FEATURES
 source
 1..1959
 /organism="Nicotiana tabacum"
 /mol_type="mRNA"
 /db_xref="taxon:4097"
 /cissue_type="leaf"
 94..1626
 /codon_start=1
 /product="myo-inositol 1-phosphate synthase"
 /protein_id="BA95788.1"
 /db_xref="GI:8096266"
 /translation="MPIENFKYSESPIVKTSEHSISHSYDYOTTELVEEKNQYOMTV
 KRKYVKEKTDVHPKLGVMVKGWGNNGSTLTGVIANRSGISATKDKYQOANVF
 GSILQASTIRVSSFNCEIYAAPKSLIPVNPDDVYFGWDSGNLADAMARAYFD
 IDLQKLRPMESVPLPGITPDPIAQAQSSRANNVITGTRKEQIDQIIRKREFKE
 NKKVKKVVLWNTANTERSNVVGVLDNTMENLFASVDRNEAISPSTLYAIACTLENV

PFINSGPONTFVPGILIDIAIKRNTLIGDDFKSQTKKSVLVDFLNGAIKPKPSIVS
 YNHLGNDGNMISAPOTERSKIEISKNVNDVMSNALLDPBPHGVHVVIVKPYV
 GDSKRAMEYTESIEIFMGKNTIVLANTDSLLAIPILDLVLAELSTRIOLKXAGE
 GKPSFPHVATILSVLTKAPLVPGTIPVNVMSKORAMLENTILACVGLABENNMILE
 YK"
 1959
 /note="15 a nucleotides"
 polyA_site
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,1e-188 Length: 1959
 Score: 2477.00 Matches: 472
 Percent Similarity: 96.67% Conservative: 21
 Best Local Similarity: 92.55% Mismatches: 17
 Query Match: 94.11% Indels: 0
 DB: 8 Gaps: 0
 US-10-718-952-2 (1-510) x AB009881 (1-1959)
 QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValysTYrThrGluThrGluIle 20
 DB 94 ATGTTATTGAGAACTTTAAGGTTGAGAGCCCAACGTTAAGTACACCGAAAGTGAGATT 153
 QY 21 GlnSerValTYrAsnTYrGluThrThrGluLeuValHisGluAsnArgAsnGlyTYrTYr 40
 DB 154 CACTCTGTCATGATTATCAACCACTGAGTTAGTTCAATGAGGAGAAAATGGGACTTAC 213
 QY 41 GlnTPriLevalLysProLysSerValLysTYrGluPheLysThrAsnIleHisValPro 60
 DB 214 CAATGAGCTGCACGCTTAAGACTGTCAATATGAGTTCAAGACATGATTCATTTCC 273
 QY 61 LysLeuGluValMetLeuValGlyTYrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
 DB 274 AATTAGGGGTTATGCTTGTGATGGGATGGAAACATAGTTCAACCTTGACCGGTGAT 333
 QY 81 ValIleAlaAsnArgGluGlyIleSerThrPalatThrLysAspLysIleGlnGlnAlaAsn 100
 DB 334 GTTATTGCTAACAGAGAAAGAAATTTCAATGGGCAACAAAGATMAAGTGCACCAACCAAT 393
 QY 101 TYrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGlu 120
 DB 394 TACTTTGGCTCTTCACTCAGGCTTCTACTATTCAGATTTGGGCTTTTCAATGAGAAAG 453
 QY 121 IleTYrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
 DB 454 ATCTATGCTCCATTTAAAGCTCTCTTCAATGTCATCCAGATGAGTATGTTGGA 513
 QY 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
 DB 514 GGATGGACATCAGCGGATGATTTGGACAGATGCGCAAGGGCTPAAGGATTTGAT 573
 QY 161 IleAspLeuGlnLysGlnLeuArgProTYrMetGlnSerMetLeuProLeuProGlyIle 180
 DB 574 ATTTGATCTCAAAACAGATGAGGCTCTACATGAAATCTATGCTCCACTACTGCTATC 633
 QY 181 TYrAspProAspPheIleAlaAlaAsnGlnGluIleArgAlaAsnAsnValIleLysGly 200
 DB 634 TATGATTCGATTTTCATTCGCTCTAACCAAGGTCACGCGCCCAACAGTGTATCAAGGA 693
 QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
 DB 694 ACCAAGAAAGAAACAAATGATCAATTCATTAAGATATTAGGAGATTAAAGAGAAAGAC 753
 QY 221 LysValAspLysValValIleuThrThrAlaAsnThrGluArgTYrSerAsnLeuVal 240
 DB 754 AAAGTGAACAAGGCTGATGCTTGTGACCTGCTAACACTGAAGATACAGTATGTGGT 813
 QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu 260
 DB 814 GTTGACCTTAATGATACATATGAAATCTCTTGTCTGTGAGACAGAAATGAACCTGAA 873
 QY 261 IleSerProSerThrLeuTYrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280

Db 874 ATATCTCTTCACATTTGTATGATGCTGATTCCTGAAATGTCCTTTATTAAT 933
 Qy 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaGAsnThr 300
 Db 934 GGAAAGCCCAAGAACACCTTTGTGCCAGGCTCATGATTTGGCCATCAAGAGAACACA 993
 Qy 301 LeuIleGlyValAspAspPheLeuSerGlyValThrIleMetLeuSerValIleLeuValAsp 320
 Db 994 CTGATGCTGGTGTATGATCACTTTAAGAGTGTCAACCAAAATGAGTCAGTGGTGGAT 1053
 Qy 321 PheLeuValGlyValGlyIleLeuSerProThrSerIleValSerTyrAsnIleLeuGlyAsn 340
 Db 1054 TTCCTTGTTGGAGCTGGTATTAAGCCAAATCATGATGTGACCTCAACACCTTTGGGTAC 1113
 Qy 341 AsnAspGlyMetCAsnLeuSerAlaProGlnThrPheArgSerIleGlyIleSerIleSer 360
 Db 1114 AAGGAGGAGATGATGATGCTGCCCCCTCAAACTTCGGCTCAAAAGAGATCTCGAAAGT 1173
 Qy 361 AsnValIleAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlyIlePro 380
 Db 1174 AATGTTGTTGATGATGATGTTTCAAGCAATGCCATTTCTTATAGAGCTGGAGAGACCTT 1233
 Qy 381 AspIleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 400
 Db 1234 GACCATGTTGTTGATTAAGTATGTCATATGTGGAGACAGCAAAAGGCAATGGAT 1293
 Qy 401 GluTyrThrSerGluIlePheMetGlyIleValSerThrIleValIleuIleAsnThrCys 420
 Db 1294 GAGTACACATCTGAGATTTTCATGGGAGAAACACCATGTTTGGCAACATACCTTGT 1353
 Qy 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleLeuValIleGlyLeuSer 440
 Db 1354 GAGGATTCCTTTTGGCTGCTCCAAATTAATCTTGATTTGGTCTTCTTGAGTACCTAGT 1413
 Qy 441 ThrArgIleGluPheValIleGluValIleGluValIleGluValIleGluValIleGluVal 460
 Db 1414 ACCCGATTCAGCTCAAGCTGAGAGGAGGAGGATTAAGTTCACCTTCCACCCCGTGGCT 1473
 Qy 461 ThrIleLeuSerTyrIleThrIleValIleProLeuValIleProGlyIleThrProValIleAsn 480
 Db 1474 ACTATCTCAAGCTTACCTTACCAAGGCTCTCTGTGTACACCAAGTACACCAAGTGTGAAT 1533
 Qy 481 AlaLeuSerIleGlnArgIleMetLeuGluValIleMetArgIleCysValIleGlyLeuAla 500
 Db 1534 GCACTCTCAAGAGAGGAGCAATGCTTGAAGAACATTTGAGGGCTGTGGTGAATGCA 1593
 Qy 501 ProGluAsnAsnMetIleLeuGluTyrIleVal 510
 Db 1594 CCAGAGAACACATGATTCGGAATACAAA 1623
 RESULT 6
 AB032073 1931 bp mRNA linear PLN 08-SEP-1999
 LOCUS Nicotiana paniculata NpINP51 mRNA for myo-inositol-1-phosphate
 DEFINITION synthase, complete cds.
 ACCESSION AB032073.1 GI:5834499
 VERSION AB032073.1
 KEYWORDS myo-inositol-1-phosphate synthase.
 SOURCE Nicotiana paniculata
 ORGANISM Nicotiana paniculata
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Nicotiana.
 REFERENCE Hashimoto, A., Yamada, S. and Komori, T.
 Myo-inositol-1-phosphate synthase
 Published Only in Database (1999)
 JOURNAL 2 (bases 1 to 1931)
 AUTHORS Hashimoto, A., Yamada, S. and Komori, T.
 TITLE Direct Submission
 SUBMITTED (01-SEP-1999) Akiko Hashimoto, Japan Tobacco Inc., Plant
 Breeding and Genetic Research Laboratory, 700 Higashibara, Iwate,
 Toyoda-cho, Shizuoka 438-0802, Japan

(E-mail: Akiko.Hashimoto@pbgrl.jti.co.jp, Tel.: +81-538-32-7116,
 Fax: +81-538-33-6046)
 FEATURES
 source
 location/Qualifiers
 1..1931
 /organism="Nicotiana paniculata"
 /mol_type="mRNA"
 /db_xref="taxon:62141"
 1..1931
 /gene="NpINP51"
 92..1624
 /gene="NpINP51"
 /codon_start=1
 /product="myo-inositol-1-phosphate synthase"
 /protein_id="BA84084.1"
 /db_xref="GI:5834500"
 /translation="MPIENPKVESPNVKTSEHSIVDYQTTLVHDEKNGYQWTV
 KPTVYKERTVHVVKLVMLVGMGNGNSTLTGVIVANREGISWATQKVOAAYF
 GSLTQASTIRVGSFNGEELIYAPPKSLIPVNPDDVYVGGMISDMKLADMAKAFD
 ILOKOLRPYMSMVPRLGIYDPDLFAAGSRANVITGDTKEQIDQIIKIDREFE
 KKKVDVVLTANTRYSNVVYGLDNTMENLPAVDREAEISPTLVAICTLENY
 PTNPSQNTFVPPGLDLAIKNNTLIGDDPFGSKYKMSVLDPVVGAIKPTLSVS
 YNHLGNNDQMLNAPQTFRSKEISKNVVDWVSNAILYEPGEHDPVWVIRKVPY
 GDSKRAMDEVTSEIFWGGKNTIVLHNTCEDSLIAPILIDVLVLAELSTRIOKAGE
 GKPSFHPVATILSYLTTRKAPLVDPCTPVNVALSKQRAMLENILRACVGLAPENNMLE
 YK"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,72e-188 Length: 1931
 Score: 2472.00 Matches: 470
 Percent Similarity: 96.86% Conservative: 24
 Best Local Similarity: 92.16% Mismatches: 16
 Query Match: 93.92% Indels: 0
 DB: 8 Gaps: 0

US-10-718-952-2 (1-510) x AB032073 (1-1931)

Qy 1 MetPheIleGluAsnPheLeuValGluCysProAsnValIleGlyThrGluThrGluIle 20
 Db 92 ATGTTATTTAGAGAAATTTTAAAGTTGAGAGCCCCCAAGGTTAAAGTACCGAAATGAAATT 151
 Qy 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValIleGluAsnArgAsnGlyTyrTyr 40
 Db 152 CACTCTGTCTATGATTAATCAACCACTGAGTTAGTTCAAGATGAGAAATGGGACATAT 211
 Qy 41 GlnTyrIleValIleAsnTyrGluThrThrGluLeuValIleGluAsnArgAsnGlyTyr 60
 Db 212 CAATGACCGTCAAGCTTAAGACTGTCAATATGATTAAGTCAAGCTGATGTTATGCC 271
 Qy 61 LysLeuGlyValMetLeuValGlyTyrGlyAsnAsnGlySerThrLeuThrGlyGly 80
 Db 272 AATTTAGGGGTTAATGCTTGTGGATGGGTGGAACCAATGTTCAACCTTGACCGGTGGT 331
 Qy 81 ValIleAlaAsnArgGluGlyIleSerTyrAlaThrIleAspLeuIleGlnGlnAlaAsn 100
 Db 332 GTTATGCTTAACAGAAATTTATGATGGCCACCAAGATTAAGGTGCAACAGCAAT 391
 Qy 101 TyrPheGlySerLeuThrGlnIleAsnAlaIleArgValIleGlySerPheGlnGlyGlu 120
 Db 392 TACTTTGGCTCTTACTCAAGCTTCTACTATTGAGTTGGGCTTTCATAGAGAAAG 451
 Qy 121 IleTyrAlaProPheLeuSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
 Db 452 ATCTATGCTCCATTTTAAAGCTCTCTTCCAAATGTCATATCCAGATGACATGATGTTGA 511
 Qy 141 GlyTyrAspIleSerAsnMetCAsnLeuAlaAspAlaMetAlaArgAlaIleValPheAsp 160
 Db 512 GAGTGGACATCAAGCACATGAATTTAGAGATGCGATGCCAGGCTAAAGATTTGAT 571
 Qy 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
 Db 572 ATTGATCTCAAAAGAGAGTTGAGGCCCTTACATGAAATCATATGATGCCACCTCTGTTATC 631

```

QY 181 TyAspProAspPheIleAlaAaengingluArgAlaAsnValIleIysGly 200
DB 632 TATGACCCGATTTATCTGCTGCTAAACAAAGGTCACGGCCAAACAGTATCAAGGA 691
QY 201 ThrIleGlnGluIleValGlnGlnIleIleValAspIleValAlaPheIleGluAlaThr 220
DB 692 ACCAAGAAAGAACAAATGATCAATCAATTAAGATTAAGAGAGTTAAGAAAAGAAC 751
QY 221 LysValAspIleValValIleValIleuThrAlaAsnThrGluArgTyrSerAsnIleuVal 240
DB 752 AAAGTGCACAGAGTGTATGATATTTGACCTGCTAACACAGTACAGTATGCGTGT 811
QY 241 ValGlyLeuAsnAspThrMetGluIleuIleuAlaValAspArgAsnGluIleuVal 260
DB 812 GTTGGACTTAAATGACCTATGAGAAACCTCTTGTCTGTGAGACGAATAGAGCTGA 871
QY 261 IleserProSerThrIleuThrAlaIleAlaCysValMetGluAsnValIleuValIleu 280
DB 872 AATATCTCTTCCACTTGTATGCTATTTGCTGATCTTGAAGAAATGTGCTTTTAT 931
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspIleuAlaIleAlaArgAsnThr 300
DB 932 GGAAGCCCCACAAACCTTTGTCCAGGCTTCATTTGCGCCATCAAGAAAGACACA 991
QY 301 LeuIleGlyAspAspPheIleuSerGlyGlnThrIleuMetIleuSerValIleuValAsp 320
DB 992 TTGATTTGGTGGTATGACCTTAAAGAGTGTGCAACCAAAATGAGTCACTGCTGTGAT 1051
QY 321 PheIleuValGlyAlaGlyIleIleuProThrSerIleValSerTyrAsnHisIleuGlyAsn 340
DB 1052 TTCCTTGTGGAGCTGGATTAAGCCACATCAATTTAGCTACCAACATTTGGTAA 1111
QY 341 AsnAspGlyMetAspIleuSerAlaProGlnThrPheArgSerIleuIleuSerIleuSer 360
DB 1112 AATGATGAAATGATATGCTGCTCCCTCAAACTTCGCTCAAGAGAGTCTCGAAAGT 1171
QY 361 AsnValIleuAspAspMetValAsnSerAsnAlaIleuIleuTyrGluProGlyGluHisPro 380
DB 1172 AATGTTGTTGATGACATGCTTCAAGCAATGCCATCTTTATGAGCTGAGAGCACCT 1231
QY 381 AspHisValValIleuIleuIleuTyrValProTyrValGlyAspSerIleuArgAlaMetAsp 400
DB 1232 GACCAATGTTGTTGATTAAGATATGCGATATGCGAGAGACAGCAAGAGGCGCAATGAT 1291
QY 401 GlyTyrThrSerGlnIlePheMetGlyIleuSerThrIleValIleuHisAsnThrCys 420
DB 1292 GAGTACACATCTGAGATTTTTCATGGGGGAAAGAACACATTTGTTGACATATCTGT 1351
QY 421 GluAspSerIleuIleuAlaProIleIleuAspIleuValIleuIleuIleuIleuSer 440
DB 1352 GAGGATTCACCTTTAGCTGCTCCAAATTTATTTGATTTGCTCTTCTGCTGAACTCAGT 1411
QY 441 ThrArgIleGluPheIleuValGluIleuGlnGlyIleuPheHisSerPheHisProValAla 460
DB 1412 ACCCGATTCAGCTCAAACTGAAAGAGAGGATTAATTCACCTCTTCCACCCCGGTGT 1471
QY 461 ThrIleIleuSerTyrIleuThrIleuValIleuValIleuValIleuValIleuValIleu 480
DB 1472 ACTATCTCACTGCTTACCAAGAGCTCTCTGCTGTCACCAAGTACACAGTGGTGAAT 1531
QY 481 AlaIleuSerIleuGlnArgAlaMetIleuGluAsnIleuMetArgAlaCysValGlyIleuAla 500
DB 1532 GCACTCTCAAGAGAGGCGCAATGCTTGAAGACATTTGAGGCGCTTGTGTGGAAGTGA 1591
QY 501 ProGluAsnAspMetIleuGluTyrIleuVal 510
DB 1592 CCAGAGAACACATGATCTGGAATACAAA 1621
RESULT 7
LOCUS E27176 1950 bp DNA linear PAT 18-JUN-2001
DEFINITION Novel INPS gene derived from nicotiana.
ACCESSION E27176

```

```

VERSION E27176.1 GI:13026394
KEYWORDS JP 1999187879-A/1.
SOURCE Nicotiana paniculata
ORGANISM Nicotiana paniculata
REFERENCE 1 (bases 1 to 1950)
AUTHORS Shigehiro, Y. and Toshiyuki, K.
TITLE Novel INPS gene derived from nicotiana
JOURNAL Patent: JP 1999187879-A 1 13-JUL-1999;
JAPAN TOBACCO INC
COMMENT
OS Nicotiana paniculata
PN JP 1999187879-A/1
PD 13-JUL-1999
PE 26-DEC-1997 JP 1997359773
PR SHIGEHIRO YAMADA, TOSHIYUKI KOMORI
PI C12N15/09/C12N5/10, C12N15/09, C12N1:91, C12N15/00, C12N5/00,
PC C12N15/00, C12R1:91
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 92..1703.
FEATURES
source 1..1950
location/Qualifiers
/organism="Nicotiana paniculata"
/mol_type="genomic DNA"
/db_xref="taxon:62141"
ORIGIN
Alignment Scores:
Pred. No.: 2,75e-188 Length: 1950
Score: 2472.00 Matches: 470
Percent Similarity: 96.86% Conservative: 24
Best Local Similarity: 92.16% Mismatches: 16
Query Match: 93.92% Indels: 0
Gaps: 0
US-10-718-952-2 (1-510) x E27176 (1-1950)
QY 1 MetPheIleGluAsnPheIleValGluCysProAsnValIleTyrThrGluThrGluIle 20
DB 92 ATGTTATTTGAAATTTTAAGTTAGAGCCCAACGTTAAGTACACCGAAGTGAAT 151
QY 21 GlnSerValIleTyrAsnTyrGluThrThrGluIleuValHisGluAsnArgAsnGlyThrTyr 40
DB 152 CACTCTGCTATGATTAATCAACCACTGAGTTAGTTGATGATGAGAAAAATGGACATAT 211
QY 41 GlnTPrIleValIleProIleuSerValIleTyrGluIleuValHisGluAsnArgAsnGlyThr 60
DB 212 CAATGACCGCTCAACCGTCAAGCTGCAAAATATGATGTTCAAGACTGATGTCATGTTCC 271
QY 61 LysIleuGlyValMetIleuValGlyTyrGlyIleuAsnGlnGlySerThrIleuThrGlyGly 80
DB 272 AAATTAGGGGTATATCTTGTGATGGGGGTGAAACAAATGCTTCAACCTTGAACCGGTGT 331
QY 81 ValIleAlaAsnArgGlnGlyIleSerTPrAlaThrIleAspIleValIleGlnGlnAlaAsn 100
DB 332 GTTATTTGCTAACAGAGAGAAATTTCAAGGCCACCAAGATAGAGTGCACAAAGCCAAAT 391
QY 101 TyrPheGlySerIleuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlnGlu 120
DB 392 TACTTTGGCTCTTACTCAGGCTTCTATTCAGAGTTGGGTCTTTCAATGAGAGAGAG 451
QY 121 IleTyrAlaProPheIleuSerIleuProMetValAsnProAspAspIleValIleGly 140
DB 452 ATCTATGCTCCATTTAAAGCTCTCTTCCATAGTCAATCCAGATGAGCTAGTGTGGA 511
QY 141 GlyTPrAspIleSerAspMetAsnIleuAlaAspAlaMetAlaArgAlaIleValIlePheAsp 160
DB 512 GGATGGACATCAGCAGCATGAAATTTAGCAGATGCCAGTGGCCAGGCGCTAAGGTTATTTGAT 571

```

```

QY 11eAspLeuGlnIleuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 572 ATTGATCTCAAAAGAGATTGAGCCCTCATGGAATCTATGTCCTCCACTCCCTGGATC 631
QY 181 TTTAspProAspPheIleAlaIleAsnGlnGlnGluArgAlaAsnValIleYsgIy 200
DB 632 TATGACCTCGATTCTGCTGCTAACCAAGGGCTCAGTCCCAACGATGATCAAGGA 691
QY 201 ThrIleGlnGlnIleGlnIleIleIleYsgIleYsgIleYsgIleYsgIleYsgIle 220
DB 692 ACCAAGAGAGAACAAATGATTCATAATCATTAAGGATATTAGGAGATTTAAGAAAAGAAC 751
QY 221 LysValAspLysValValIleuTyrThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 752 AAAGTGAACAAGGTGATGATATTGTGAGCTGCTAACACTGAAAGATACGATATGCGTT 811
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaValAspArgAsnGluIleGlu 260
DB 812 GTTGGACTTAATGACACTATGAAAACCTCTTGTGCTGTGAGCAAGAAATGAACTGAA 871
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 872 ATATCTCTTCCACTTGTATGCTATGTGCTGATTCGTAATAATGCGCTTTTATTAAT 931
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 932 GGAAGCCCCAGAACACCTTTGTCAGGCTCATTTGTCCTCATCAAGAAAGAACACA 991
QY 301 LeuIleGlyIleAspAspPheIleYsgIleGlnThrLysMetLysSerValLeuValAsp 320
DB 992 TTGATGATGGTGATGATCTTTAAGATGGTCAAAACCAATGAAAGTCAGTGGTGGAT 1051
QY 321 PheLeuValGlyAlaIleYleYsProThrSerIleValSerTyrAsnIleGluIleAsn 340
DB 1052 TTCCTTGTGGAGCTGTATTAAGCCACATCAATGTGCTACCAACCACTTGGGTAAAC 1111
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleYsgIleSerLysSer 360
DB 1112 AATGATGAAATGATATGTCCTGCTCCCTCAAACTTTCGGTCMAAGAGATCTCGAAAAGT 1171
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluIlePro 380
DB 1172 AATGTTGTGATGACATGCTTCAACCAATGCAATCTTATGAGCTGAGAGACCCCT 1231
QY 381 AspHisValValIleYsTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1232 GACCAATGTTGTGATTAAGTATGCAATATGTCGAGAGACGAAAGGGCAATGAGAT 1291
QY 401 GlnTyrThrSerGlnIlePheMetGlyIleYsSerThrIleValLeuHisAsnThrCys 420
DB 1292 GAGTACACATCTGAGATTTTCATGGGGGAAAGAACACCAATGTTTTCACATTAATCTGT 1351
QY 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1352 GAGGATTCACATTTAGCTGCTCCAAATTATATTGATTTGCTCTTCTGCTGAACTCAAT 1411
QY 441 ThrArgIleGluPheIleValIleGluAsnGluIleYsPheHisSerPheHisProValAla 460
DB 1412 ACCCGCATTCACCTCAAAAGCTGAAGGAGGGGTAACTTCACCTTCCACCCCGGGCT 1471
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB 1472 ACTATCTCAAGCTACTTACCAAGGCTCTCTGGATCCACGAGTACACACAGGGTGAAT 1531
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1532 GCACTCTCAAGCAGAGGGCAATGCTTGAACAATATGAGGGCTTGTGTGGAATTTGCA 1591
QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
DB 1592 CCAAGAAACAAATGATCTGGAATTAACAA 1621

```

RESULT 8

```

BT013505
LOCUS      BT013505                1959 bp.  mRNA.  linear.  PLN 11-MAY-2004
DEFINITION Lycopereisicon esculentum clone 132193F, mRNA sequence.
ACCESSION  BT013505
VERSION    BT013505.1
KEYWORDS   FLI CDNA.
SOURCE     Lycopereisicon esculentum (tomato)
ORGANISM   Lycopereisicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
            asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopereisicon.
REFERENCE  1 (bases 1 to 1959)
            Kirkness, E.F., Wang, W. and Vazeille, A.
AUTHORS    Submitted (11-MAY-2004) The Inetlicute for Genomic Research, 9712
            Medical Center Drive, Rockville, MD 20850, USA
JOURNAL    Location/Qualifiers
FEATURES   1..1959
            source
            /organism="Lycopereisicon esculentum"
            /mol_type="mRNA"
            /db_xref="taxon:4081"
            /clone="132193F"
            /issue_type="fruit"
            /note="TXBFS1"
ORIGIN
Alignment Scores:
Pred. No.:      4,81e-188
Score:          2469.00
Percent Similarity: 96.27%
Best Local Similarity: 92.55%
Query Match:    93.81%
DB:             8
Gaps:           0
US-10-718-952-2 (1-510) x BT013505 (1-1959)
QY 1 MetPheIleGluAspPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 77 ATGTTATTATGAATAATTTTAAGGGAGAACCCCAATGTAAGATATTTGAGAAATGAAT 136
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValIleGluAsnArgAsnGlyThrTyr 40
DB 137 CATTCGTGTATGATTATTAACACACAGAGCTTGTTCAAGAGAGAGAAATGGAATCTTAT 196
QY 41 GlnTyrIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
DB 197 CAATGATGTTTAAGCTTAACATGTCAAATATGAATTTAAACTGATACCCATGCGCA 256
QY 61 LysLeuGlyValMetLeuValGlyTyrProGlyIleAsnAsnGlySerThrLeuThrGlyIle 80
DB 257 AATATGGGGGTATGCTGTTGTGATGGGAGGAGAAACAAATGTTCAACATTTGAGAGGT 316
QY 81 ValIleAlaAsnArgGluIleYsSerTyrAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 317 GTTATTGGCAATCGAAGAGATTTTCATGCGACAGAAAGAAAAGTGCACAGCCAAAT 376
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluIle 120
DB 377 TATTTTGGGTCTCTTAATCTAGGCAATCAACCATTCAGTGGGTCTTTCAATGGCGAAGAG 436
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 437 ATCTATGACCCCTTCAAAAGCTCTCTCCCATGCTCAACCCAGACGATGATGATTTGGA 496
QY 141 GlnTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 497 GGATGGACATTAAGCAACATGAATTTGGCAAGATGCTATGTCAGGGCTTAAGTTTCGA 556
QY 161 IleAspLeuGlnIleGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 557 GTTGATCTGCAAAACAGCTGAGGCTTACAGAGAAATCATGATGTTCCCTTCCTGATATC 616
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGlnIleGluArgAlaAsnValIleYsgIy 200

```



```

Db      617 TATGACCCCTGACTTCATTCGGCGCTAAACAAGACGCGCAACAACCTGATCAAGA 676
Qy      201 ThrTlrsGIngluGlnGlnGlnIleIleLysAspIleLysAlaPheLysGluIaThr 220
Db      677 ACCAAGAAGAACCAAGTTGAACAATGTTAAAGATTAATTAAGAGATTCAAGAGAAAC 736
Qy      221 LysValAspLysValValIleuThrPThrIaAsnThrGluuArgLysSerAsnLeuVal 240
Db      737 AAGTAGACACAGATAGTGTTCTATGAGACTGCACACCGAAGGATAGATATGCGTT 796
Qy      241 ValGlyLeuAsnAspThrMetGluLeuLeuLeuAlaValAspArgAsnGluIaGlu 260
Db      797 GTTGCGCTTAATGACCACTGAGAAAACCTTTAGCTGCTGAGTAAGAAATAGGCTGA 856
Qy      261 ILeSerProSerThrLeuThrValIaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db      857 AATATCTCTTCTACATTCGATGCTATTCCTGTTGATTAATGAAAATGTCCTTTCATCAC 916
Qy      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      917 GGAAGCCCTCAABAACCTTTGTTCCAGGCTCTTATGATTTGGCCATTAAAGAAACA 976
Qy      301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db      977 TTAATTGGGTGTGATGACTTTAAGAGTGGTCAAAACCAAGATGAAGTCCGTGTGTTGAT 1036
Qy      321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisIleuGlyAsn 340
Db      1037 TTCTCTGTGTGAGCCGGATATTAAGCCACGTCATATGATGACATACATTCGGGTAC 1096
Qy      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db      1097 AATGATGAATGATATCTTCTGCTCTCAACCTTCGGTCTTAAGAGATCTCAAAAGT 1156
Qy      361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db      1157 AATGTTGTGATGACATGTTGCTAGCAACGCCATCTCTAATGAATCTGCGACACCT 1216
Qy      381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db      1217 GACCATGTTGTGATGACAAATGTTCCATATGTTGACACGCAAGAGCGCAATGAT 1276
Qy      401 GluTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
Db      1277 GAGTACATGTCAGAGATTTTTCATGGCGGAAAGAGACATATGTTATGCAACAACCTTGT 1336
Qy      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db      1337 GAGGACTCTCTTTGGCACGCTCCAAATTAATTCGATTTGCTCTCTCGTCAACTCAC 1396
Qy      441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db      1397 ACTCGCATTCAGCTCAAGCTGAAGGGAGGAGAAATTCATCTCTTCCACCTGTGGCT 1456
Qy      461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db      1457 ACTATTCTCAGCTACTCTCAACAAGGCTCCCTGGTATGCCACAGTACTCTCAAGTGGAT 1516
Qy      481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db      1517 GCCCTTTCAAGCAGAGCGCAATGCTTGAGAAATATATGAGAGCTTGTGTGTTGGCA 1576
Qy      501 ProGluAsnAspMetIleLeuGluTyrLys 510
Db      1577 CCAGAGAACACATGATTTTGGAAATCAAA 1606

```

```

RESULT 9      AF284065      1845 bp      mRNA      linear      PIN 30-JUL-2003
LOCUS      AF284065
DEFINITION      Sesamum indicum myo-inositol 1-phosphate synthase mRNA, complete
            cds.
ACCESSION      AF284065

```

```

VERSION      AF284065.1 GI:9858815
KEYWORDS      .
SOURCE      Sesamum indicum (sesame)
ORGANISM      Sesamum indicum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots;
            asterids; lamids; Lamiales; Pedalaceae; Sesamum.
REFERENCE      1 (bases 1 to 1845)
AUTHORS      Pyun,J.A., Jin,U.H., Lee,J.W., Yi,Y.B., Hyung,N.I., Kang,M.H.,
            Pyun,J.H., Suh,M.C., Kang,C.W., Seo,H.Y., Lee,S.W. and Chung,C.H.
TITLE      Isolation and characterization of a myo-inositol 1-phosphate
            synthase cDNA from developing sesame (Sesamum indicum L.) seeds:
            functional and differential expression, and salt-induced
            transcription during germination
JOURNAL      Planta 216 (5), 874-880 (2003)
MEDLINE      22511750
PUBMED      12624775
REFERENCE      2 (bases 1 to 1845)
AUTHORS      Jin,U.-H. and Chung,C.-H.
TITLE      Direct Submission
JOURNAL      Submitted (01-JUL-2000) Division of Biotechnology, Faculty of Life
            Science and Resources, Dong-A University, Ha-Dan-dong, Sa-Gu-Gu,,
            Pusan City, Pusan 604-714, South Korea
FEATURES
            source
            1..1845
            /organism="Sesamum indicum"
            /mol_type="mRNA"
            /db_xref="taxon:4182"
            /tissue_type="seeds"
            /note="authority: Sesamum indicum L."
            51..1583
            /codon_start=1
            /product="myo-inositol 1-phosphate synthase"
            /protein_id="AA01148.1"
            /db_xref="GI:9858816"
            /translation="MFISFVSESNVYKTEGHSYVYETTELVHSENGTYOMIV
            KPTVYKFKIDHYVPLGLVMVGKGNNGSTLTGVIANBGISMAKDKYQQANF
            GSLTQASSIRVGSFNGEBITVAPFKSLIPVNPDDVDFGWMISNNMLADWAGRAVL
            IDLOKOLRPYHEHWPLPGIYDPPIANQSRANNVKGRKEQVOQI IKDMRPFKE
            QNKVDKVVLTWANTERSNVVGLNDPESLMSVENEAEISPSLYAIACPEVN
            PRINSPONTFVPLIDAIQNRSLIGDDPFKSGQTKRKSVDLVLGAGIKPTIS
            YNHLNNDGNMLSAQTRSKSISNVVDMMVNSGILYEGEHPDHIIVYKYVYV
            GDSKAMDEYSEIFMGKSTIVLHNTCEDSLIAPILDLVLAEISTRIQKAGE
            GKFSFHHVATILSYLTKAPLVPPGTPVNALSKQRAWLENILRA CVGLAPENNMLE
            YK"
ORIGIN
Alignment Scores:
Pred. No.:      8,53e-187      Length:      1845
Score:      2453.00      Matches:      470
Percent Similarity:      96.08%      Conservative:      20
Best Local Similarity:      92.16%      Mismatches:      20
Query Match:      93.20%      Indels:      0
DB:      Gaps:      0
US-10-718-952-2 (1-510) x AF284065 (1-1845)
Qy      1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
Db      51 ATGTTCAATCGAAGAGCTTCAAGGTTAGAGACCCCAATGTCAAGTACCAAGAGGTAGAT 110
Qy      21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db      111 CACTCGTCTACCACTACGAGACTACAGAGCTTGTTCATGAGAGAGAGAAACGGCACCTAC 170
Qy      41 GlnTTPILValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
Db      171 CAATGATTTGTCACCAAGCCTGCAAAATCGAGTTCAAAACATGATATGTATGCTCC 230
Qy      61 LysLeuGlyValMetLeuValGlyTTPGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db      231 AATATTAGGGGTCATCCTGTGTGATGGAGGAGCAACATGTTCACCTCACTGCGCGGT 290

```


QY 81 Val11ealAaenArgGlugly11eSerTrpAlaThrLysAspLys11eGlnGlnAlaen 100
 DB 291 GTCATTGCAACAGGAGAGGATTTCAATGGCAACAAAGATBAGGTGCAAGGCCAAT 350
 QY 101 TyrPheglYserLeuThrglnAlaSerAla1eArgValGlySerPheglnglyGln 120
 DB 351 TATTTGGGTCACTGACACAGGATCTCAATCAGATTGGTTCTTCAATGAGAAAG 410
 QY 121 11eTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAl1eValPheglY 140
 DB 411 ATCTATGCCCCCTTCAAAAGCCCTTCTCCATGGTCAACCCAGATACGATAGTGGG 470
 QY 141 GlyTPAsp11eSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
 DB 471 GCAATGGACATACGCAATATGAATTTGGCTGATGCCATGGCCAGGCGCAAGGTTAGAT 530
 QY 161 11eAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProglY11e 180
 DB 531 ATTGATCTCCAGAGCAGCTCAGGCCCTTCAATGACACATGGTCCCATCTCCCTGGAATC 590
 QY 181 TyrAspProAspPhe11eAlaAlaenGlnGlnLysValAsnAsnVal11eLysGly 200
 DB 591 TACGATCCCTGATTTACGCGCCATCAGGATCAGTCCCAACACGATTAAGA 650
 QY 201 ThrLysGlnGlnValGlnGln11e11eLysAsp11eLysValPheLysGlnAlaThr 220
 DB 651 ACCAAGAAAGAACATTTCAACGATCATCAAGACATGAGGACCTTTAAGAGCAAAAT 710
 QY 221 LysValAspLysValValLeuTPrThrAlaAsnThrglnArgYserAsnLeuVal 240
 DB 711 AAGGTGACAGAGTGTGCTAGTGTGTCAGACAGCCAAACAGAAAGATACGATATGTTGTC 770
 QY 241 ValGlyLeuAsnAspThrMetGluLeuLeuAlaValAspArgAsnGlnValGlu 260
 DB 771 GTTGGCTTAAAGATACACCCCAAGCCGATGCGCTGATGAGAGATAGGCGGAG 830
 QY 261 11eSerProSerThrLeuTyrAla11eAlaCysValMetGluAsnValProPhe11eAsn 280
 DB 831 AATATCTCTTCAACCTTGTATGACATAGCTGTGTTTGAAGAAATGTCGTTCAATAT 890
 QY 281 GlySerProGlnAsnThrPheValProGlyLeu11eAspLeuAla11eAlaArgAsnThr 300
 DB 891 GGCAGGCCCAAAATATCTTTGTTCAGGGTTGATTTGGCGATTCAGAGGAACACT 950
 QY 301 Leu11eGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
 DB 951 TTGATTTGGTGAAGATGATTTCAAGAGCGGTCAAAACCAAGATGAATCAGTGTGAT 1010
 QY 321 PheLeuValGlyAlaGly11eLysProThrSer11eValSerTyrAsn11eGlnGlyAsn 340
 DB 1011 TTCTTGTGAGAGCTGGATCAAGCCCAAGCTGATGTGATATATCATCTGGGAAAC 1070
 QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlu11eSerLysSer 360
 DB 1071 AACGACGGGATATATCTATCAGACACCCCAACTTCCGCTCCAAAGAGTTCTTAAGAAC 1130
 QY 361 AsnValValAspAspMetValAsnSerAsnAla11eLeuTyrGluProGlyGln11ePro 380
 DB 1131 AATGTGTGAGACGACATGCTTCCAGCAACGCGCATCTTTATGAGCTGTGATAACCCCC 1190
 QY 381 AspH1eValVal11eLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
 DB 1191 GACCATATTTGTTCATCAAGATGATCCGTAATGTGGGAGACGACAGACATAGGAGAC 1250
 QY 401 GlyTyrThrSerGln11ePheMetGlyGlyLysSerThr11eValLeuH1eAsnThrCys 420
 DB 1251 GAGTACACTTCCGAGATATTATCAGGAGGAGAAAGATACATAGTTCTGCAACACGTGT 1310
 QY 421 GluAspSerLeuLeuAlaAlaPro11e11eLeuAspLeuValLeuLeuAlaGluLeuSer 440
 DB 1311 GAGGACTCTCTTGTGCTGCTCCCATCATCTGGATTTGTGCTCTCTACTACTAAGCTACAC 1370
 QY 441 ThrArg11eGluPheLysAlaGluAsnGlnGlyLysPheH1eSerPheH1eProValAla 460

DB 1371 ACTCGATTCACTCAAAAGCCAGAGAGGCGCAATTCATTCTTCCACCGGTGGCA 1430
 QY 461 Thr11eLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
 DB 1431 ACCATCTCAGCTACCTTAACCAAGCCCTTGTATCTTCCAGGACACCGGTGTCAC 1490
 QY 481 AlaLeuSerLysGlnArgAlaMetLeuGlnAsn11eMetArgAlaCysValGlyLeuAla 500
 DB 1491 GCACTGTCTAACAGAGGTGCATGCTCGAGAACATCTTGAGGCGTGTGGATTTGGCT 1550
 QY 501 ProGluAsnAsnMet11eLeuGlnTyrLys 510
 DB 1551 CCAAGAAACATATGATTTGGAAATACAG 1580
 RESULT 10
 BT013759 1954 bp mRNA linear PLN 11-MAY-2004
 LOCUS
 DEFINITION Lycopersicon esculentum clone 132642F, mRNA sequence.
 ACCESSION BT013759
 VERSION BT013759.1 GI:47105174
 KEYWORDS FLI_CDNA.
 SOURCE Lycopersicon esculentum (tomato)
 ORGANISM Lycopersicon esculentum
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 1954)
 AUTHORS Kirnesh, E.F., Wang, W. and Vazellie, A.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAY-2004) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA
 FEATURES
 source
 1..1954
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /db_xref="taxon:4081"
 /clone="132642F"
 /tissue_type="fruit"
 /note="TZXB41"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,21e-185 Length: 1954
 Score: 2439.00 Matches: 463
 Percent Similarity: 95.88% Conservative: 26
 Best Local Similarity: 90.78% Mismatches: 21
 Query Match: 92.67% Indels: 0
 DB: 8 Gaps: 0
 US-10-718-952-2 (1-510) x BT013759 (1-1954)
 QY 1 MetPhe11eGlnAsnPheLysValGluCysProAsnValLysTyrThrglnThrgln 20
 DB 107 ATGTTATTAGAACTTTAAGGTTGAGAGCCCAATGTTAATACACAGAGTGAATT 166
 QY 21 GlnSerValTyrAsnTyrGluThrThrglnLeuValH1eGluAsnArgAsnGlyThrTyr 40
 DB 167 CACTCGTGTATGATTAATCAACACAGTGAAGTTGTTCAATGAGAAAGAAAAATGGACATTT 226
 QY 41 GlnTPr11eValLysProLysSerValLysTyrGlnPheLysThrAsn11eH1eValPro 60
 DB 227 CAATGACTGTAAACCCAAAACCTGTCAAAATTTGAATTCAAAACGTGATGTATGTTCCA 286
 QY 61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrglyGly 80
 DB 287 AAATTAGAGATTATCTTGTGATGGGTGGAAGACATGTTCAAGCTTGAACCGAGGT 346
 QY 81 Val11eAlaAsnArgGlnGly11eSerTrpAlaThrLysAspLys11eGlnGlnAlaAsn 100
 DB 347 GTTATTTGCTAATAGAGAAAGAAATTTCAATGGGCTACCAAGATTAAGTACAGCAACCAAT 406
 QY 101 TyrPheglYserLeuThrglnAlaSerAla1eArgValGlySerPheglnglyGln 120

```

Db      407 TACTTGGTGTCTTACACAGGCTCTACTATTGAGTGGGCTTTCAATGAGAGGAG 466
Qy      121 IleyrAlaProPhelyserleuLeuProMetValAsnProAspAlaPheGly 140
Db      467 AACTATGCTCCCTTTAAGAGCACTACTCTCCATGTCATCAAGATGATGATTTTGG 526
Qy      141 GYTPAPAPILeserAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db      527 GGATGGACATACACACATGATTTGGACATGCCATGGCAAGACTTAAAGTTTGGAC 586
Qy      161 ILeasPLeuGlnLysGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyIle 180
Db      587 ATTGATCTTCAATGCAACTGAGGCGGTACATGAAATCATGATGATCCCATGCTGGTATC 646
Qy      181 TYRAspProAspPheIleAlaAlaengInGluArgAlaAsnAsnValIleLysGly 200
Db      647 TATGACCTGATTTCACTGCTGCAACCAAGGCTCAGGTGTAACATGTCATCAAGGA 706
Qy      201 ThrLysGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
Db      707 ACCAAGAAAGAACCAATTATCATCAATTAATAAGATATTAGGAGCTTTAAGAGAACAA 766
Qy      221 LysValAspLysValValIleuTyrPheAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db      767 AAAGTGAACAGGATGGTGTCTCTGTCGACCTGCCAACATGAAAGATACAGCAGTGGCT 826
Qy      241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgGlnGluArgIle 260
Db      827 GTTGGCTTTAATGATACCATGAGAAACCTTTGCTGTGTGATAGAAATAGAGGTGA 886
Qy      261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db      887 AATCTCTCTTCCACCTGTGATGCTATGCTTGTGATCTTGAAATGATACCTTCATCAT 946
Qy      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      947 GGAAAGCCCAAAATATCTTTGTCGCCAGGCTCATGATTTGGCCATCAAGAGAACACT 1006
Qy      301 LeuIleGlyGlyAspAspPheLysSerGlyInThrLysMetLysSerValLeuValAsp 320
Db      1007 TTGATTTGGTGGTGTGATCTTTAAGAGTGTGTCAAACCAATGAAAGTCACTGCTGTGAT 1066
Qy      321 PheLeuValGlyValGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db      1067 TTCCTTTGGTGGAGCTGTATTAAGCCAAATCATGATGATGATCAACCATTTGGGTAA 1126
Qy      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnLysSerLysSer 360
Db      1127 AATGATGAAATGATCTGTCTGCCCCCTCAAACTTTCCGTTCAAGAGAGATCTCGAAGAAC 1186
Qy      361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyGlnHisPro 380
Db      1187 AACGTTGTGATGACACGATTTCTAGTAAACGCATCTTTATAGAGCTGAGAGCACTCTCT 1246
Qy      381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db      1247 GACCACTGTTGTTGTATTAAGTACATCATATGTCGAGATGACAAAGGCGAAATGAT 1306
Qy      401 GlyTyrThrSerGlnLysPheMetGlyLysSerThrIleValLeuHisAsnThrCys 420
Db      1307 GAGTACACATCAGAGATTTTCACTGAGTGAAGGAACACATGATGATTCACAATACCTGT 1366
Qy      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db      1367 GAGGATTTCTTTTGGCTCTCCAAATTAATCTTGATTTAGTGTCTGTGCTGAACTCAAT 1426
Qy      441 ThrArgIleGluPheLysAlaGluAsnGlnGlyLysPheHisSerPheHisProValAla 460
Db      1427 ACCCGATTCAAGTCAAGCTGAAGAGAGGGTAATTCATCTCTTCATCCCGTGGCG 1486
Qy      461 ThrIleLeuSerTyrIleuThrLysAlaProLeuValProProGlyThrProValValAsn 480

```

```

Db      1487 ACCATCTCAGTTATCTTACCAAGGCTCCTTGTGATCACACAGGTACACCACTGTGAT 1546
Qy      481 AlaSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db      1547 GCCCTTTCAAGACAGAGGCAATGCTCGTGAACATTAATTGAGGCTTGTGTAAGTGGCA 1606
Qy      501 ProGluAsnAsnMetIleLeuGlnTyrLys 510
Db      1607 CCAGAGAACCAATCATGATCTTGGAATACAA 1636

```

```

RESULT 11
AF120146      1538 bp      mRNA      linear      PLN 23-APR-1999
LOCUS
DEFINITION
Triticum aestivum cultivar Fielder myo-inositol 1-phosphate
synthase (MIPS) mRNA, complete cds.
ACCESSION
AF120146
VERSION
AF120146.1 GI:4589061
KEYWORDS
SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 1538)
Hussein,A., Yan,W., Bock,C., Baga,M., Chibbar,R. and Georges,F.
CDNA clone for myo-inositol 1-phosphate synthase from wheat
Unpublished
JOURNAL
2 (bases 1 to 1538)
Hussein,A., Yan,W., Bock,C., Baga,M., Chibbar,R. and Georges,F.
Direct Submission
Submitted (14-JAN-1999) National Research Council Canada, Plant
Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N
0W9, Canada

```

```

FEATURES
source
1..1538
location/Qualifiers
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Fielder"
/db_xref="taxon:4565"
1..1538
/gene="MIPS"
1..1533
/gene="MIPS"
/EC_number="5.5.1.4"
/function="converts glucose 6-phosphate into myo-inositol
1-phosphate"
/note="cytosolic enzyme"
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAD26330.1"
/db_xref="GI:4589062"
/translaction="MIFSPKVESPNVYKTENEHSHVYDETTEVVENRNGTYQWV
KRTYKVPKIDTRPKGLWLVGKNGGSLTGVANKEGISMATGDKYOANVF
GSLTQASSIRGSIVGEBEIVAPFKSLPMVNPDDVDFGMDISDNMLADNARAYLD
IDLQKLRPMYEMWPLPGIYDPIANQGSRAVSIGTKKEOVDOIHKMREFE
KMKVDKLVLTANTERSYDVVGVNDTMENTLASEKNESESISSTIYALACVEGI
PINGSPOINTFVPGILELISKNCLIGDDFKSGTCKSVLADFLVGAGIKPTISV
YNHLNNGGMLNLSAQOTRSKEISKSNVDVVAAGLIEBGEHPDHVVYKVPYV
GDSKRAMBYSEIFMGKNTVMNTCEDSLAAPIILDVLIAELSTRIOFKAGE
GKFSHFHPATILSYLTATPLVPAQTPIVVALSKORAMLENTLRACVGLAPENNMLE
YK"

```

```

CDS
gene
1..1538

```

ORIGIN

```

Alignment Scores:
Pred. No.:      5,2e-182      Length:      1538
Score:          2392.00      Matches:      453
Percent Similarity: 95.88%      Conservative: 36
Best Local Similarity: 88.82%      Mismatches: 21
Query Match:      90.88%      Indels:      0
DB:              8          Gaps:      0

```

US-10-718-952-2 (1-510) x AF120146 (1-1538)

```

QY 1 MetPhe11eGluAenPhe1yVal1GluCysProAenVal1ySerThyTrhGluThrGlu11e 20
Db 1 ATGTTTCATGAGAGGCTTCAAAAGTTGAGAGCCCAAGCTGAACTACACAGAAACGAGATC 60
QY 21 GlnSerVal1yAenThyTrhGluThrGluLeuVal1h1eGluAenAryaenGlyThr1y 40
Db 61 CATTCGTGTGAGATTACGAAACACAGGAGTGTCTCATAGAAACGCTAACGGCACTTAC 120
QY 41 GlnTrp11eVal1ySerPolySerVal1ySerGluPhe1yThrAen11eH1sValPro 60
Db 121 CATGTGGCTGTCAAGCCCAAAACGGTCAATACGATTCAAAAACAGCATCCGCTGCC 180
QY 61 LysLeuG1yValMetLeuVal1G1yTrpG1yG1yAenAenG1ySerThrLeuThrG1y 80
Db 181 AACCTGAGGCTTATGCTCGTTGAGTGGGAGAAACATGATCAACCTTAACCGCTGGA 240
QY 81 Val11eAlaAenAryG1y11eSerTrpAlaThrLysAenP1yS11eGlnGlnAlaAen 100
Db 241 GTGATGCCCAAGAAAGAGGATCTGTGGCGACTTAAGACAAAGTGCACAAAGCAAC 300
QY 101 TyrPheG1ySerLeuThrGlnAlaSerAla11eAryG1yG1ySerPheGlnG1yGlu 120
Db 301 TACTTTGGCTCCCTCACTCAAGCTTCTCGATCCGGGTGGATCTTAACAATGGTGAAG 360
QY 121 11eTyrAlaProPhe1ySerLeuLeuProMetVal1aenProAenP11eValPheG1y 140
Db 361 ATCTAGGCTCTTTCAGAGCCCTCTCCCATGTGAACCTCAAGATGTGTGTGGT 420
QY 141 G1yTrpAsp11eSerAenMetAenLeuAlaAenAlaSerAla11eAryG1yA1yValPheAsp 160
Db 421 GGTGGGACATTAAGCAGCATGAACCTTAAGCAGACGCTATGCTAGAGAAAGTCTTGAC 480
QY 161 11eAspLeuGln1yGlnLeuAryPro1yMetG1ySerMetLeuProLeuProG1y11e 180
Db 481 ATTGATCTCCAGAAACAGCTCAGCGCTCATGAGACATGTGTCCACACTCCCTGGATC 540
QY 181 TyrAspProAenPhe11eAlaAenGlnGluG1yAenG1yAenAenVal11eLysG1y 200
Db 541 TACGACCTGACTTCACTCGTGCAGATCAAGGGTCAAGCTCTAATAGCCGATCAAAAGT 600
QY 201 ThrLysGlnGluGlnVal1GlnGln11eLysAsp11eLysAlaPhe1ySerG1yAlaThr 220
Db 601 ACCAATAAAGAAACAATCGACGACGATCATTAAGACATGAGGAGTTTAAGAAAGAAC 660
QY 221 LysVal1aenP1yVal1aenVal1eUtrP1yThrAlaAenThrG1yUtrG1ySerAenLeuVal 240
Db 661 AAGGTGATTAAGCTGTGTCTGTGTGAGACGCAACACAGAGGTTACAGCGATGTGTC 720
QY 241 Val1G1yLeuAenAenP1yMetG1yAenLeuLeuAlaVal1aenAryaenG1yAlaGlu 260
Db 721 GTAGGGCTTAACGACGATGAGAAATCTAATGGCTTGTGAAGAAACGAGTGTGAG 780
QY 261 11eSerProSerThrLeuThrAla11eAlaCysVal1MetG1yAenVal1ProPhe11eAen 280
Db 781 ACTGCTCTTCAACGCTATATGCGATTCGTGTCTTGAAGGCACTCTTCTTCAATAT 840
QY 281 G1ySerProGlnAenThrPheVal1ProG1yLeu11eAspLeuAla11eAlaAryaenThr 300
Db 841 GGGAGTCCCGACAGAACGTTTGTCTTGAGACTCATGAGCTGTGCTATATCAAGAACTGT 900
QY 301 Leu11eG1yG1yAenAenPhe1ySerG1yGlnThrLysMetLysSerVal1LeuValAen 320
Db 901 TTGATCGGTGGGATACTTCAAGAGTGTGCAACATTAAGTGAAGTCTGTGTGTGAT 960
QY 321 PheLeuVal1G1yValaG1y11eLysProThrSer11eValSerTyrAenH1sLeuG1yAen 340
Db 961 TTCCTTGTGGAGCTGCACTCAAGCTTCACTTCAATGTGATGCTTAATCATTTGGGAAAC 1020
QY 341 AenAenP1yMetAenLeuSerAlaProGlnThrPheArySerLysG1y11eSerLysSer 360
Db 1021 AATGACGGATGAACCTTCAAGCGCTCAAGACTTTAGATCTTAAGAGATCTCAAGAGAG 1080
QY 361 AenVal1aenAenPhe1yValaenSerAenAla11eLeuTyrG1yProG1yG1yH1sPro 380

```

```

Db 1081 AATGTTGTGACACACATGTTGCTATTAACGATCTCTTTAGACCTGTGAACATCT 1140
QY 381 AspH1sVal1aenVal11eLysTyrVal1ProTyrVal1G1yAenSerLysAryAlaMetAen 400
Db 1141 GACCATGTTGTGTCAATCAAGATGTCTTATGTGGGATGTAGTAAGAGACCATAGAC 1200
QY 401 G1yTyrThrSerGln11ePheMetG1yG1ySerThr11eValLeuH1sAenThrCys 420
Db 1201 GAGTACACCTCGGAGATTAATTCAGGAGGAGAAATACATGTTATGATCAATACCTTGT 1260
QY 421 GluAenSerLeuLeuAlaPro11eLysAenP1yLeuAenVal1LeuLeuA1yLeuAenSer 440
Db 1261 GAGATTCCTTTTGGCTGTCTGTCCATCATCTTGATCTTGTCTTCTGCTGAGCTCAGC 1320
QY 441 ThrAryG1eGluPheLysAlaGluAenG1yG1yLysPheH1sSerPheH1sProVal1a 460
Db 1321 ACCAATCCAAATTAAGACTGAAGGAGGAGGAGGAGATTCATTTCTTCCACCCAGTAGCA 1380
QY 461 Thr11eLeuSerTyrLeuThrLysAlaProLeuVal1ProProG1yThrProVal1ValAen 480
Db 1381 ACCATCACTAGTTACTTCAAAAGGACCTCTGTACACAGCCGAGACACCTGTGTGAAC 1440
QY 481 AlaLeuSerLysGlnAryAlaMetLeuGluAen11eMetAryAlaCysVal1G1yLeuAla 500
Db 1441 GCTTATCTTAAGCAGAGGCTATGCTGAGAACATCTTCAAGGCTGTGGCTTGGC 1500
QY 501 ProGluAenAenMet11eLeuGluTyrLys 510
Db 1501 CCAGAGAACACATGATCTTGGATACAG 1530

```

RESULT 12

AF120147
LOCUS
DEFINITION
Triticum aestivum cultivar Taber myo-inositol 1-phosphate synthase (MIPS) mRNA, complete cds.

ACCESSION
AF120147.1 GI:4589063

VERSION
AF120147.1

KEYWORDS
Triticum aestivum (bread wheat)

SOURCE
ORGANISM
Triticum aestivum

REFERENCE
AUTHORS
Hussein, A., Yan, W., Bock, C., Baga, M., Chibbar, R. and Georges, F.
TITLE
CDNA clone for myo-inositol 1-phosphate synthase from wheat
JOURNAL
Unpublished

REFERENCE
AUTHORS
Hussein, A., Yan, W., Bock, C., Baga, M., Chibbar, R. and Georges, F.
TITLE
Direct SubMISSION
SUBMITTED (14-JAN-1999) National Research Council Canada, Plant Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Canada

FEATURES
source
1..1538
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultiyar="taber"
/db_xref="taxon:4565"
1..1538
/gene="MIPS"
1..1533
/gene="MIPS"
/EC_number="5.5.1.4"
/function="converts glucose 6-phosphate into myo-inositol 1-phosphate"
/note="cytosolic enzyme"
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAD26331.1"
/db_xref="GI:4589064"
/translation="MFIESFKVSEENVKTEENHSVDYETTEVVHNRNGTYQMVV"

Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Canada
Location/Qualifiers
1..1538
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Biggar"
/db_xref="taxon:4565"
1..1538
/gene="MIPS"
1..1533
/gene="MIPS"
/EC_number="5.5.1.4"
/function="converts glucose 6-phosphate into myo-inositol 1-phosphate"
/note="cytosolic enzyme"
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAD26332.1"
/db_xref="GI:4589066"
/translation="MPTSPKVESPVVKTTENEHSVYDETTEVHNKNGTYQNVKRTVYDKTDTRVPLKLVMLVWGNNSTLTAGVANKGI SWATDXYQANVF GSITQASIRVGSYNGEIEYAPFKSLPMVNPDDVFGMDISDMNLADAMARVLD ILQKOLRYMEHNVPLPGIYDPDPIAANGSRANSVIGTKKEVDIOIIXOMREKCKKVKDLVYLVMTANTERSDVVVGINDMTMENISVKKESRISPTLYAIAICVREGI PRINSPOPTFVPGLEIAISKNCLEGGDFKSGQTKMSVLDPLVGAGIRPTSYV YNHLGNDMDNISAPQTFRSKEISKNVDDVANSGLFPEGEHDDHVAVIKYEVY GSKRAMDEYTSIEIPWGKNITVMHNTCDSLLAPlIDLVLMELSTRIOFKEGE GKFHSFHPATILSYLTAKAPlVPACTPVVNAISKORAMENILRACVGLAPENMILSYR"
ORIGIN
Alignment Scores:
Pred. No.: 5..2e-182 Length: 1538
Score: 2392.00 Matches: 453
Percent Similarity: 95.88 Conservative: 36
Best Local Similarity: 86.82 Mismatches: 21
Query Match: 90.88 Indels: 0
Gaps: 0
DB: 8
US-10-718-952-2 (1-510) x AF120148 (1-1538)
QY 1 MetPheIleGluAsnPhelValGluCysProAsnValIleThrGluThrGluIle 20
DB 1 ATGTTTCATGAGAGCTTCAAAGTTGAGAGCCCAAGCTAAAGTACACAGAGATC 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHleGluAsnArgAsnGlyThrTyr 40
DB 61 CATTCGTGTTACGATTACGAACACGAGGTCGTCTCATGAGAACCTTAACCGCATTAAC 120
QY 41 GlnThrIleValIleProlSerValIleTyrGluPheIleThrAsnIleHleValPro 60
DB 121 CAATGGGTGTCAGAGCCCAAAACCGTCAATACGATTCAAAACACAGACATCGCGTCCC 180
QY 61 LysIleuGluValMetIleuValGluTyrGluYAsnAsnGlySerThrIleuThrGly 80
DB 181 AAGCTAGGCGTTATGCTGTGATGAGGAGAAACAATGAGTCAACCTTAACCGCTGGA 240
QY 81 ValIleAlaAsnArgGluGluIleSerTyrAlaThrIleYAsnArgIleGlnGlnAlaAsn 100
DB 241 GTGATCGCCAAACAAGAGATCTGTGGCCATTAAGACAAAGTGCACAGCAAGAAC 300
QY 101 TyrPheGlySerIleuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTTGGCTCCCTCACTCAAGCTTCTTCGATCCGGGTGGATCTTAACAATGGTGAAG 360
QY 121 IleTyrAlaProPheIleSerIleuLeuProMetValAsnProAsnArgIleValPheGly 140
DB 361 ATCTAGCGCTCTTCAAAGAGCTCTCCCATGATGTAACCCGTGACATGTGTGGTTGGT 420
QY 141 GlyTyrAsnIleSerAsnMetAsnLeuAlaAsnAlaMetAlaArgAlaIleValPheAsp 160
DB 421 GGTGGGACATAAGCAATGAATTAAGACGCTATGGCTTAGACAAAGATCTTGGAC 480

QY 161 IleAspLeuGlnIleGluIleuArgProTyrMetGluSerMetLeuProIleProGlyIle 180
DB 481 ATTGATCTCCAAACAAGCTCAGGCCGTACATGAGACACATGATCTCCCTGGATATC 540
QY 181 TyrAspProAsnPhelIleAlaIleAsnGlnGluGluArgAlaAsnAsnValIleTyrGly 200
DB 541 TACGACCTGATCTTCACTGCTCGAATCAAGGTCACGTCCTTAATAGCTGATCAAGGT 600
QY 201 ThrIleGlnGluGlnValGlnGlnIleIleTyrAspIleIleValAlaPheIleGluAlaThr 220
DB 601 ACCAAGAAAGAACCAAGTCCAGCAGATCATCAAAAGACATAGAGGAGATTTAAGAGAAAGAAC 660
QY 221 LysValAspIleValValIleuThrPThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAGGTGATTAAGCTGTTCTGTGACAGGCAACACAGACGTTACAGCATGTGTGTC 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsnGluAlaGlu 260
DB 721 GTAGGCTTACGACACATGAGAAATCTATTGGCTTCTGTGAGAAACAGATCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATCTGCTTCTACGCTATATGCGATGCTGTGTGTTCTTGAAGGATTCCTTTCATTAT 840
QY 281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGGAGTCCCAAGAACACGTTTGTCTTGACATCATCGACTGCTATATACAGAACTGT 900
QY 301 LeuIleGlyIleAspAsnPhelSerGlyIleThrIleMetLysSerValLeuValAsp 320
DB 901 TTGATCGGTGGGAGTACCTCAAGAGTGTCAAGCTTACATGAAGTCTGTTTGGTTGAT 960
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHleuGluAsn 340
DB 961 TTCCTTGTGGAGCTGGATCAAGGCTACTTGATTTGAGCTATATATCATTTTGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleGluIleSerLysSer 360
DB 1021 AATGACGGCATAACTCTCAGCGCTCAGACTTTTGAATCTTAAGAGATCTCAAGAAC 1080
QY 361 AsnValIleAspAsnMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHlePro 380
DB 1081 AATGTTGTTGACAGATGAGTGTCTAGTACGTTATCTTTCAGCTGTGTAAACATCTC 1140
QY 381 AspHleValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCATGTTGTGTCATCAAGATATGTTCTTATGTTGGGATAGTAAAGAACCATGTGAC 1200
QY 401 GluTyrThrSerGluIlePheMetGlyLysSerThrIleValIleuHleAsnThrCys 420
DB 1201 GAGTACACCTCCGAGATATTTCTGAGGAGGAAGATACCATGTGTTATGCATTAACA 1260
QY 421 GluAspSerLeuLeuAlaProIleIleLeuAspLeuValIleuLeuAlaGluLeuSer 440
DB 1261 GAGGATCTCTTTGTGCTCAATCAATCTTGATCTTGTCTTCCTGCTAGCTCAAGC 1320
QY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHleSerPheHleProValAla 460
DB 1321 ACCAGATATTCATTAAGAGCTGAAGGAGGAGGAGATTTCATTTCCACCAAGTAGA 1380
QY 461 ThrIleLeuSerTyrLeuThrIleAlaProLeuValProProGlyThrProValValAsn 480
DB 1381 ACCATATCTACAGTTACCTCAAAAGGACACTTGTGAACAGCGGAGACCTGTGTGTAAC 1440
QY 481 AlaLeuSerLysGluArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1441 GCTTATCTTAAGCAGAGGCTATGCTGAGAAACATCTCAAGGCTGTGGTGGCTTGG 1500
QY 501 ProGluAsnAsnMetIleuGluTyrLys 510
DB 1501 CCAAGAACACATGATCTTGAATTAACAG 1530

RESULT 14
AF433879 1986 bp mRNA linear PIN 07-NOV-2001
LOCUS
DEFINITION
Suaeda maritima subsp. salsa myo-inositol-1-phosphate synthase
(INPS) mRNA, complete cds.
ACCESSION
AF433879 GI:16755888
KEYWORDS
Suaeda salsa
ORGANISM
Suaeda salsa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Suaeda.
REFERENCE
AUTHORS
Wang, J. P., Wang, P. P., Sun, Y. F., Zhao, Y. X. and Zhang, H.
TITLE
Direct Submission
JOURNAL
Submitted (15-OCT-2001) The Biology Department of Shandong Normal
University, Key Laboratory of Plant Stress Research, No.88, Wenhua
East Road, Jinan, Shandong Province 250014, China
Location/Qualifiers
1. 1986
FEATURES
source
/organism="Suaeda salsa"
/mol_type="mRNA"
/sub_species="salsa"
/db_xref="taxon:126914"
1. 1986
/gene="INPS"
140. 1672
/gene="INPS"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAI28131.1"
/db_xref="GI:16755889"
/translation="MPIENPKYSEPVNVTENEHISVYVYETTELAVHNRKSGYOWI
VAPKIVKIQTHVFKLGVMLVWGKNGCTLTGCVIANRBSGIMATKDIQANY
FESLTQASSIRVGSFNGEETIYAPFKSLPMVNPDDI VFGWPISDLNADATRAKVL
DIDLQOLAPMEHWPLPGIYDPFIANODSRANNIIKGTKEQVEQVIDIEFK
EKXKVDVVLWNTATERYSDVVGINDTENLMAVERSEI SPSTLYAACIOED
VPIFGSPQNTFPGILIELAIKNCLIGDDRSQOTKMSVLVDFIVAGIKPISIV
SYNHGNNDCMNLSPQTPRSKEISKSNVDMVANSGLYVRGSHPDHVVYIKVYV
VSDSKRAMEYTSSEIFMGSKNTIVLHNTEDSLAAPTILDLVLAESTRIEFAET
EDKFTFHPVATILSYLTKAPLVPPTPVNALSKORAMLEKRAVCVGLAPNNMILE
YK"

ORIGIN

Alignment Scores:

Pred. No.:	7,12e-182	Length:	1986
Score:	2392.00	Matches:	460
Percent Similarity:	94.72%	Conservative:	24
Best Local Similarity:	90.02%	Mismatches:	25
Query Match:	90.88%	Indels:	2
DB:	8	Gaps:	2

US-10-718-952-2 (1-510) x AF433879 (1-1986)

QY 1 MacPhe1leG1uAsnPhelYsValG1uCySProAsnValYerThrG1uThrG1u1e 20
Db 140 A1GTTTATGAAATTTCAAGTGAAGGACCCAAAGCTTAACCTAAGTAAATTT 199
QY 21 G1nSeVAlYrAsnTYrG1uThrThrG1uLeuValH1eG1uAsnA1rGAsnG1YThr 39
Db 200 CATTGAGTTTACAAATATGAACCACTGAAATAGTTTACAGAAATGCAAGGCTGTGCT 259
QY 40 TYrG1nTPr1leValYsPrO1ySseRValYsTYrG1uPhelYsThrAn1leH1sVal 59
Db 260 TATCATATGATCGTCAAACTTAACCAAGCTTCAATATACCAATTCAAAATCTCAATGTT 319
QY 60 PrO1yS1eUeG1YValMeTleuValG1YTPG1YGLYAsnAsnG1YSeThr1eUThrG1Y 79
Db 320 CCCAAACTTGGGGTATGTTGGTGGGATGGGGTGGCAACACGTTGCCACCTCACCGGT 379
QY 80 G1YVal11eAlAAsnA1rG1uG1Y1eSeTTPAlaThr1YsA1p1Ys11eG1nG1nAla 99

Db 380 GGTTCATCGCCATTCGGAGGGTATATCATGGGCACAAAGACAAATTCACACACT 439
QY 100 AsnTYrPheG1YSeTleuThrG1nAlaSeRAla1eArGValG1YSeRPhelG1nG1Y 119
Db 440 AACATATTTGGATCCCTCACTCAGGCTTCTTATTCGTGTGGTCTTTAATGGAAG 499
QY 120 G1u1eTYrAlaPrOphelYsSeR1eUeUPrOveTValAaNPPrOAspA1leVal1he 139
Db 500 GAGATTTATGCTCTTTTAAGAGCTTCTCTCATGATGTAACCCAGATGATATAGTGT 559
QY 140 G1YGLYTPAAsP1leSeRAsnMeTAsn1eUAlaAsPAlaMeTAlaArGAl1eYsAl1he 159
Db 560 GGGGGCTGGACATTAAGACTTGAACCTAGACAGCGCATGACGAGGCAAGGATGTTG 619
QY 160 AsP1leAsP1eUeG1n1eUeYrPrO1YrMeT1eUeG1SeRMeT1eUePrO1Y 179
Db 620 GACATTTGATTTGCAAAAGACAGTTGACCTTACATGAGACATGTTCCCTCTCTGT 679
QY 180 11eTYrAsPPrOAsPPhel1eAla1aAsnG1nG1uArGAlaAsnAsnVal11eYs 199
Db 680 ATCTATGACCTGATTTGCTCTCTCAACCAAGATTCGTCGCAATATACATCTTAAG 739
QY 200 G1YThr1YrG1nG1uG1nValG1nG1n11e11eYsAaP11eYsAlaPhelYsG1u1a 219
Db 740 GGCATTAAGAAAGACAAAGTGGAGCAAGTCATTAAGACATCAGGAATTAAGAGAG 799
QY 220 Th1YsValAsP1YsValVal1eUThrPrTh1aAsnThrG1uArGTYSeRAsn1eU 239
Db 800 AACAAAGTGGACAAAGTGTGTACTCTTGACAGCAAAACAGAGAGTACATGATGTG 859
QY 240 ValValG1YLeuAsnAsPThrMeT1eUAsn1eUeUAla1aValaAsPArGAsnG1u1a 259
Db 860 GTGGTGGGACTCAATGACACCAAGAACTTATGGCAGCTTGGAAGAATAGTCT 919
QY 260 G1u1eSeRPrOSeRTh1eUThrAla11eAlaCySValMeT1eUAsnValPrOhe11e 279
Db 920 GAGATTTTCCATTCACATTTGTATGCTATGCTTGACCAAGAAAGTCCCTTCAATC 979
QY 280 AsnG1YSeRPrO1aAsnThrPhelValPrO1YLeu1eAsP1eUAl1eAlaArGAsn 299
Db 980 AACGATAGTCTCAGAACACCTTTTCTCTGTCTCATTTGAACCTGCCATTAAGAAAT 1039
QY 300 Th1eU11eG1Y1aSPAsPPhelYsSeR1Yg1nThr1YsMeT1eUSeRVal1eUVal 319
Db 1040 TGTTCATTTGAGAGAGATGACTTCAAGAGTGTGACCTAAGATGAATATCAGTCTCTT 1099
QY 320 AsPPhel1eUValG1YAlaG1Y11eYsPrOThrSeR11eValSeTYrAsnH1eUeG1Y 339
Db 1100 GACTTCCTTGTGGGGCTGGATTAAGCCAACTCAATGTTGAGCTACCAACCACTTGGGA 1159
QY 340 AsnAsnAsP1YMeTAsn1eUSeRAlaPrO1nThrPhelArGSeR1Yg1U1eSeR1Ys 359
Db 1160 AACAAATGATGGATAACTTTCAGACCTCAGACCTTCAGGTCATTAAGAAATTTCAAA 1219
QY 360 SeRAsnVal1aAsPAsPMeTValaSnSeRAsnAla11eUeTYrG1uPrO1Yg1uH1s 379
Db 1220 AGTATATGTTGATGATGATGTTGCTAGCAACGCTATCTGTATGAGCGCTGTGAACAC 1279
QY 380 PrOAsPPh1sVal1aVal11eYsTYrValPrO1YrValG1YAsPSeR1YArGAlaMeT 399
Db 1280 CCTGATCATGTTGTGTATCATCAAGTATGTCATATGTCGAGAGACGAAAGAGACTATG 1339
QY 400 AsP1YrThrSeR1Y11ePhelMeT1Yg1YsSeRTh1eVal11eU11eAsnThr 419
Db 1340 GATGAGTACATCTGAGATATTCATGGGTGATTAAGAACACATGCTTTTGCAACAACC 1399
QY 420 CyS1uAsPSeR1eUeUAlaAlaPrO11e1eUAsP1eUVal11eUeUAlaG1u1eU 439
Db 1400 TGTGAGACTCTCTTGTGGCGCTCCAAATCATCTTGAAGCTTGCTCTTGGCGAGCTT 1459
QY 440 SeThr1ArG11eG1uPhelYsAlaG1uAsnG1uG1Y1ePh1eH1sSeRPhel1ArOVal 459
Db 1460 AGCACCCGATTAATTCAAAGCCGAAACCTGAGGACAAAGTTCCACACCTTCATCTCAAGT 1519


```
Db 1262 AAGAGCAACGTGTCATGATGTTGCCAGCAACGGCATCCTCATGAGCCGGCGAA 1321
Qy 379 H|S|P|R|A|S|P|H|S|V|A|I|V|A|I|I|L|E|Y|T|Y|V|A|I|P|R|O|T|Y|V|A|I|G|Y|A|S|P|S|E|L|Y|S|A|R|G|A|I|A 398
Db 1322 C|A|C|C|T|G|A|T|C|A|T|G|T|T|G|T|G|T|C|A|G|T|A|T|G|T|G|C|C|A|T|A|T|G|T|G|G|A|C|A|C|A|G|A|G|A|G|C|C 1381
Qy 399 M|E|T|A|S|P|G|I|U|T|Y|R|T|H|S|E|R|G|I|U|I|E|P|H|E|T|G|I|Y|G|I|Y|S|E|R|T|H|I|I|E|V|A|I|L|E|U|H|I|S|A|N 418
Db 1382 A|T|G|A|T|G|A|G|T|A|C|C|T|C|G|A|A|A|T|A|T|T|A|T|G|G|G|G|G|A|A|C|G|A|C|A|C|T|A|T|G|T|C|A|T|G|C|A|C|A|C 1441
Qy 419 T|H|C|Y|S|G|I|U|A|S|P|S|E|L|E|U|A|I|A|A|P|R|O|I|E|I|I|E|U|A|S|P|L|E|U|V|A|I|L|E|U|A|I|A|G|I|U 438
Db 1442 A|C|C|T|G|A|G|A|C|T|C|T|C|T|T|G|G|C|T|G|C|C|C|C|A|T|C|A|T|C|T|G|A|C|T|T|G|G|T|C|C|T|G|C|T|G|A 1501
Qy 439 L|E|U|S|E|R|T|H|A|R|G|I|I|E|G|I|U|P|H|E|L|Y|S|A|I|A|G|I|U|A|N|G|I|U|G|I|Y|S|P|H|E|I|S|E|R|P|H|E|I|S|P|R|O 458
Db 1502 C|T|A|G|C|A|C|C|C|G|A|T|T|C|A|G|C|T|G|A|A|G|G|C|T|G|A|G|A|G|A|G|A|G|A|G|T|T|C|C|A|C|T|C|T|C|A|T|C|C|G 1561
Qy 459 V|A|I|A|I|A|T|H|I|I|E|U|S|E|R|T|Y|R|E|U|T|H|I|Y|S|A|I|A|P|R|O|L|E|U|V|A|I|P|R|O|F|R|O|G|I|Y|T|H|R|P|R|O|V|A|I 478
Db 1562 G|T|C|G|C|T|A|C|C|A|T|C|T|G|A|G|T|T|A|C|C|T|C|A|C|A|A|G|G|C|C|C|C|T|T|G|T|T|C|C|A|C|A|G|G|C|A|C|T|C|C|A|G|T|G 1621
Qy 479 V|A|I|A|S|N|A|I|L|E|U|S|E|R|Y|S|G|I|A|R|A|I|A|E|T|L|E|U|G|I|U|A|S|N|I|I|E|M|E|C|A|R|G|A|I|A|C|Y|S|V|A|I|G|I|Y 498
Db 1622 G|T|A|A|A|T|G|A|C|T|T|C|G|A|A|G|C|A|G|A|G|G|G|C|T|A|T|G|C|T|G|A|G|A|C|A|C|A|C|T|A|C|T|G|A|G|G|G|C|T|G|T|G|G|T 1681
Qy 499 L|E|U|A|I|A|P|R|O|G|I|U|A|S|N|A|S|N|E|C|I|I|E|L|E|U|G|I|U|T|Y|R|Y|S 510
Db 1682 T|T|G|G|C|T|C|T|G|A|G|A|C|A|C|A|T|G|A|T|C|C|T|G|A|A|T|A|C|A|A|G 1717
```

Search completed: June 7, 2005, 20:45:00
Job time : 5362 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: June 7, 2005, 17:10:41 ; Search time 633.5 Seconds
(without alignments)
4765.695 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632
Sequence: 1 MLENPKVCECPVKTETETI.....NIMRACVGLAPENMILEYK 510

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10718952/runat_06062005.173400.12840/app.query.fasta_1.1294
-DB=N Geneseq.16Dec04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-NODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10718952.0CGN_1.1.886.0runat_06062005.173400.12840 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Geneseq.16Dec04:*
2: geneseqn1908:*
3: geneseqn1908:*
4: geneseqn2000a:*
5: geneseqn2001a:*
6: geneseqn2002a:*
7: geneseqn2002a:*
8: geneseqn2003a:*
9: geneseqn2003a:*
10: geneseqn2003a:*
11: geneseqn2003a:*
12: geneseqn2004a:*
13: geneseqn2004a:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2632	100.0	1533	12	ADQ14498
2	2632	100.0	1533	13	ADQ14498
3	2632	100.0	1533	13	ADQ14498
4	2632	100.0	1760	12	ADQ14490
5	2632	100.0	1782	2	AAV62440

6	2627	99.8	1533	2	AAV62443	AAV62443 Soybean m
7	2627	99.8	1533	12	ADQ14494	ADQ14494 Mutant so
8	2627	99.8	1533	13	ADQ14494	ADQ14494 Mutant so
9	2600	98.8	1533	12	ADQ14504	ADQ14504 Wild type
10	2600	98.8	1533	13	ADQ14502	ADQ14502 Mutant so
11	2600	98.8	1533	12	ADQ14502	ADQ14502 Mutant so
12	2600	98.8	1533	13	ADQ14502	ADQ14502 Mutant so
13	2593	98.5	1533	13	ADQ14500	ADQ14500 Mutant so
14	2593	98.5	1533	13	ADQ14500	ADQ14500 Mutant so
15	2472	93.9	1530	2	AAV62443	AAV62443 Soybean m
16	2378	90.3	1533	12	ADQ14494	ADQ14494 Mutant so
17	2378	90.3	1533	12	ADQ14494	ADQ14494 Mutant so
18	2378	90.3	1533	12	ADQ14494	ADQ14494 Mutant so
19	2368	90.0	1781	4	AAV62443	AAV62443 Soybean m
20	2352.5	89.4	1536	3	AAV62443	AAV62443 Soybean m
21	2352.5	89.4	1537	3	AAV62443	AAV62443 Soybean m
22	2344	89.1	1665	2	AAV62443	AAV62443 Soybean m
23	2336	88.8	1931	2	AAV62443	AAV62443 Soybean m
24	2335	88.7	1959	4	AAV62443	AAV62443 Soybean m
25	2283	86.7	1759	6	AAV62443	AAV62443 Soybean m
26	2129.5	80.9	1719	3	AAV62443	AAV62443 Soybean m
27	2026	77.0	1536	12	ADQ14494	ADQ14494 Mutant so
28	1813	68.9	3546	2	AAV62443	AAV62443 Soybean m
29	1812	68.8	3546	2	AAV62443	AAV62443 Soybean m
30	1607	61.1	1772	4	AAV62443	AAV62443 Soybean m
31	1607	61.1	1818	4	AAV62443	AAV62443 Soybean m
32	1607	61.1	1825	6	AAV62443	AAV62443 Soybean m
33	1607	61.1	1852	4	AAV62443	AAV62443 Soybean m
34	1607	61.1	1862	4	AAV62443	AAV62443 Soybean m
35	1607	61.1	2380	13	ADQ14494	ADQ14494 Mutant so
36	1607	61.1	2380	13	ADQ14494	ADQ14494 Mutant so
37	1607	61.1	2380	13	ADQ14494	ADQ14494 Mutant so
38	1596.5	60.7	1605	8	AAV62443	AAV62443 Soybean m
39	1593	60.5	1633	4	AAV62443	AAV62443 Soybean m
40	1592.5	60.5	1605	8	AAV62443	AAV62443 Soybean m
41	1585	60.2	2280	4	AAV62443	AAV62443 Soybean m
42	1556.5	59.1	1980	12	ADQ14494	ADQ14494 Mutant so
43	1556.5	59.1	1980	12	ADQ14494	ADQ14494 Mutant so
44	1547	58.8	1991	12	ADQ14494	ADQ14494 Mutant so
45	1530	58.1	1704	8	AAV62443	AAV62443 Soybean m

ALIGNMENTS

RESULT 1	ADQ14498	ADQ14498 standard; cDNA; 1533 BP.
ID	ADQ14498	
XX	ADQ14498	
AC	ADQ14498	
DT	23-SEP-2004	(first entry)
DE	Mutant soybean myo-inositol 1-phosphate synthase cDNA #2.	
XX		
KW	Soybean, myo-inositol 1-phosphate synthase; gene; ss;	
KW	myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;	
KW	inosinic phosphate; mutant.	
XX		
OS	Glycine max.	
OS	Synthetic.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	1..1533
FT		/*tag= a
FT		/product= "Mutant soybean myo-inositol 1-phosphate
FT		synthase #2"
PN	US2004128713-A1.	
XX		
XX		
PD	01-JUL-2004.	
XX		
PF	21-NOV-2003; 2003US-00718952.	
XX		

PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-00506822.
PR 26-APR-1999; 98US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
DR WPI; 2004-533135/51.
XX P-PSDB; ADQ14499.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phylic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX Example 8; SEQ ID NO 9; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

XX SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,98e-253	Length:	1533
Score:	2632.00	Matches:	510
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-10-718-952-2 (1-510) x ADQ14498 (1-1533)

QY 1 MetPheIleGluAsnPhelYsValGluCysProAsnValIysTyrThrGluThrGluIle 20
DB 1 AAGTTCATCGAAGATTTTAAGGTGAGTGCTCAATGTGAAGTACCCGAGACTGAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGCGTGTACCACTACGAAACCAACGAACTGTTCACGAGAACGGAATGGCACTTAT 120
QY 41 GlnTrrIleValIysProIysSerValIysTyrGluPhelYsThrAsnIleHisValPro 60
DB 121 CAGTGAATGTCAACCAACCAATCTGTCAATACGAATTTAAACCAACATCTCACTTCTT 180
QY 61 LysLeuGluValMetLeuValGlyTrrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 181 AATTTAGGGGTATGCTTGTTGGGTGGGATGAGAAACAACGGCTCAACCTTCACCGTGGT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTrrPalaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATTGTCAACGAGGAGGCAATTCATGGGCTACAAAGACAAAGATTCACCAAGCAAT 300

QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGlu 120
DB 301 TACTTTGGCTCCCTCACCACCAAGCTTCAGACTATCCGAGTTGGGTCTCTTCCAGGAGAGAAA 360
QY 121 IleTrrAlaProPheIysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCTCACTTCMAAGAGCTGCTTCCAAATGGTTAAACCTCGAGACAAATGTGTGGG 420
QY 141 GlyTrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGATATCAGCAACAATGAACTGGCTGATGCTGATGCTGAGGCAAGAGGTCTTTGAC 480
QY 481 ATCGATTTGCAAGAACAGTGTAGGCTTACATGAAATCAATGCTTCCACTCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaAsnGlnGluValArgAlaAsnValIleLysGly 200
DB 541 TATGACCCGGATTTCTATTGCTGCCAACCAAGGAGCGTGCACMAAGTCAATCAAGGAGC 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleTrrLysAspIleLysAlaPheLysGlnAlaThr 220
DB 601 ACAAGCAAGAGCAAGTTCACAAATCATCAAGACATCAAGCGCTTTAAGAGCCACC 660
QY 221 LysValAspLysValValIleuTrrPrrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAGTGGACAAGGTGGTGTGACTGTGACTGCACTGCCCAACAGAGGTAAGTATTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaAlaValAspArgAsnGluAlaGlu 260
DB 721 GTGGGCTTTAAGCACACATGAGATCTCTTGCTGTGTGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTrrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCTCTTCCACTGTGATGCTGCTGTGTGTAAGAAATGCTTCTTCAATTAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAGAACACTTTGTACAGAGGCTGATTGATCTTGCCATGCGGAGAACACT 900
QY 301 LeuIleGlyIleAspAspPheLysSerGlyGlnTrrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTTGGGAGAGACTTCMAAGTGTCTCAGACCAAAATGAAATCTGTGTGGTTGAT 960
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 961 TTCCTTGGGGGGCTGGTATCAAGCCAACTATTAAGTCAAGTCAACCATCTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1021 AATGATGGTATGAATCTTGGCTCCCAAACTTCCGTTCCAAAGAAATCTCCAAAGAGC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1081 AACGTTGTGATGATATGCTCAACAGCAATGCAATCTCTTAAGACCCTGGTGAACATCCA 1140
QY 381 AspHisValValIleLysTrrValProTrrValGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCAATGTTGTTATTAAGTATGATGCTTACGTAAGGAGACAGCAAGAACCTGAT 1200
QY 401 GlnTrrThrSerGluIlePheMetGlyLysSerThrIleValLeuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGATATTTACGTGGTGGAAAGAACACATCTTTTGACAAACATGTC 1260
QY 421 GluAspSerLeuLeuAlaLarProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTCTTCTGCTGCTCTTATTAATCTTGAGCTTGCTCTTCTGCTGAGCTCAGC 1320
QY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAAGATCGAATTTAAAGCTGAAAATGAGGAAAATTCATCTCAATCCACGATTGCT 1380
QY 461 ThrIleLeuSerTrrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480

DB 1381 ACCATCTGAGCTACACCAAGGCTCCTCTGTTCCACCGGGTACACCAAGTGGATT 1440
QY 481 A1A1eSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1441 GCATTGTCGAAGGAGCGGTCAATGCTGGAAAACATATGAGGGCTGTGTTGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
DB 1501 CCAGAGAAATPACATGATTCTCGATCAAG 1530
RESULT 2
ADS81999
ID ADS81999 standard; cDNA; 1533 BP.
XX
AC ADS81999;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 3.
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; line 29004JP01.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /product= "myo-inositol 1-phosphate synthase"
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
PI Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
DR P-PSDB; ADS82000.
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX
PS Example 8; SEQ ID NO 9; 34p; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a

CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with a
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.98e-253 Length: 1533
Score: 2632.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Gaps: 0
DB: 13 Indels: 0
US-10-718-952-2 (1-510) x ADS81999 (1-1533)
QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 1 ATGTCATCGAAGAAATTTAAAGTTAGTGTCTTAATGTAATACCAAGACTGAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnLysTyrTyr 40
DB 61 CAGTCGGTGTACACATCAGCAAAACACCGAAGCTGTTTCACGAGAACAGAAATGGACCTAT 120
QY 41 GlnTTPlleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
DB 121 CAGTGGATTGTCAAAACCCAAATCTGCAAAATGCAATTTAAACCAACATCCATGTTCTT 180
QY 61 LysLeuGluValMetLeuValGlyTyrPGLysAsnAsnGlySerThrLeuThrGlyGly 80
DB 181 AAATTAAGGGGTAAATCTGTGGGTGGGAGGAAACAACGGCTCAACCTCACCGGTGGT 240
QY 81 ValIleAlaAsnArgGluGlyLysSerTTPAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATTGCTAACCGAGAGGGCATTTTCATGGCTTACAAGAGAACAGATTCAACAGCAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTGGCTCCCTCACCAAGCTCAGCTATCCAGTGGGTCTTCCAGGAGAGAGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ACTTATGCCCTTCATAAGAGCTGTCTTCATGTTAACCTCGAGACATTTGTGTTGGG 420
QY 141 GlyTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGGATATACGAACATGAACCTGGCTATGCTACAGGCCAAGAGTGTTTGAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGlnSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTGGCAAGACAGATGAGGCTTATACGAAATCATGCTCTCCACCTCCCGAAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 541 TATGACCCGGATTTATTTATGTCGCAACCAAGAGGCGGCCAACACGTCATCAAGGGC 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleLysAspIleLysValAlaPheLysGlnAlaThr 220
DB 601 ACAAGACCAAGGACCAAGTTCACAAATCATCAAGACATCAAGCGCTTTAAGAACCAACC 660
QY 221 LysValAspLysValValValLeuTTPThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGACCAAGGTGTGTTACTGTGACTGCAACACAGAGAGGTACAGTAATTTGGTT 720
QY 241 ValGlyLeuAsnArgTPThrMetGluAsnLeuValAlaValAspArgAsnGluAlaGlu 260

```

Db      |||||
721 GTGGGCTTAATGACCATGAGAACTCTTGCGCTGTGAGACAGAAATGAGGCTGAG 780
Qy      |||||
261 IIESerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db      |||||
781 ATTTCTCTTCCACCTTGATGATGCAATGCTGTGTATGAAAGATGTTCTTTCAATTAAT 840
Qy      |||||
281 GlycerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaIleAsnThr 300
Db      |||||
841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTGATCTTGCCATCGCGAGAACACT 900
Qy      |||||
301 LeuIleGlyGlyAspAspPheIleYserGlyGlnThrLysMetLysSerValIleuValAsp 320
Db      |||||
901 TTGATTGGGTGAGATGACTTCAAGAGTGGTCAAGACCAAAATCTGTGTGGTGTGAT 960
Qy      |||||
321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db      |||||
961 TTCTCTTGGGGGGCTGGATGCAAGCCACATCTATAGTACATTACACCATCTGGGAAAC 1020
Qy      |||||
341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyIleSerLysSer 360
Db      |||||
1021 AATGATGATGATGATCTTTCGGCTCCAAACTTTCGGTCCAGGAATCTCCAAAGACC 1080
Qy      |||||
361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db      |||||
1081 AACGTTGTTGATGATGATGCTCAACAGCAATGCCATCTCTATGAGCCTGTCACATCCA 1140
Qy      |||||
381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db      |||||
1141 GACCATGTTGTATTAAGTATGATGCTTACGTAAGGAGACAGCAAGAGCATGAT 1200
Qy      |||||
401 GluTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValIleuHisAsnThrCys 420
Db      |||||
1201 GAGTACACTTCAGAGATATTCAATGGGTGGAAGAGACACATGTTTGGACAAACATCTC 1260
Qy      |||||
421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleuLeuIleuSer 440
Db      |||||
1261 GAGGATTCCTCTTGAGCTCTCTCTATTAATCTTGAACTTGCTGCTTCTGTCAGCTCAGC 1320
Qy      |||||
441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db      |||||
1321 ACTAGATGAGATTAAAGCTGAAATGAGGAAATTCACATTCATTCACACCATGTTGCT 1380
Qy      |||||
461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValIleAsn 480
Db      |||||
1381 ACCATCTCAGCTACTCTCAACAGGCTCTCTGTTCCACCGGTATCACACAGTGGAT 1440
Qy      |||||
481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db      |||||
1441 GCATTGTCAAAGCAGCGTCAATGCTCGAATAACATTAATGAGGCTTGTGATTGGCC 1500
Qy      |||||
501 ProGluAsnAsnMetIleLeuGluTyrLys 510
Db      |||||
1501 CCAGAGATTAACATGATCTTCGAGTACAAAG 1530

```

RESULT 3
AD881993 standard; cDNA; 1533 BP.

AD881993; 18-NOV-2004 (first entry)
Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 1.
Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
Glycine max; cultivar Wye.
Key Location/Qualifiers
1..1533
CDS /tag= a

```

FT      /product= "myo-inositol 1-phosphate synthase"
XX      US2003074685-A1.
PN      17-APR-2003.
XX      11-MAR-2002; 2002US-00025003.
XX      08-APR-1997; 97US-00835751.
PR      07-APR-1998; 98WO-US006822.
XX      (HITZ/) HITZ W D.
PA      (SEBA/) SEBASTIAN S A.
XX      Hitz WD, Sebastian SA;
PI      WPI: 2004-639957/62.
DR      P-PsDB; AD881994.
XX      Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT      phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT      phytic acid and inorganic phosphate content of soybean seeds.
XX      Claim 2; SEQ ID NO 1; 34pp; English.
PS      The invention relates to an isolated nucleic acid fragment encoding a
CC      soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC      phosphate synthase having decreasing capacity for the synthesis for myo-
CC      inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC      nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC      complement, subfragment or the complement of the subfragment, operably
CC      linked to suitable regulatory sequences), where expression of native gene
CC      gene results in a decrease in expression of an endogenous or native gene
CC      encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC      comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC      acid content of less than 17 micromol/g, a seed content of raffinose plus
CC      stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC      greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC      from the plant, making a soybean plant with the heritable phenotype
CC      (comprising crossing LR33 or the plant comprising the chimeric gene with
CC      an elite soybean plant and selecting a progeny plant of the cross of
CC      crossing step that has a heritable phenotype as mentioned above), seeds
CC      of soybean plant made by the above method, a soy protein product derived
CC      from seeds of a soybean plant (homozygous for one or more gene encoding a
CC      mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC      the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC      heritable phenotype as mentioned above), and making or producing a
CC      soybean protein product derived from seeds of a soybean plant with a
CC      heritable phenotype as mentioned above. The nucleic acid is useful for
CC      altering raffinose saccharide, sucrose, phytic acid and inorganic
CC      phosphate content of soybean seeds thus leading to valuable and useful
CC      soybean products, since the presence of high concentration of raffinose
CC      oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC      when consumed by humans. The present sequence encodes a wild-type myo-
CC      inositol 1-phosphate synthase.
XX      SQ      Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

```

Alignment Scores:

Score:	1.98e-253	Length:	1533
Percent Similarity:	2632.00	Matches:	510
Best Local Similarity:	100.00%	Conservative:	0
Query Match:	100.00%	Mismatches:	0
DB:	13	Indels:	0
		Gaps:	0

US-10-718-952-2 (1-510) x AD881993 (1-1533)

```

Qy      1 MetPheIleGluAsnPhelYsValGluCysProAsnValIleYrThrGluThrGluIle 20
Db      1 ATGTTTCATGAGAAATTTTAAAGTTAGTGTCTTAATGCAATACCGAGCATGAGATT 60
Qy      21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40

```

```

Db 61 CACTCGGTGACAACTACGAAACCAACGACTTGTTACAGAAACAGAGATGACACTAT 120
Qy 41 GATTTPILLeValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
Db 121 CAGTGAATGTCACAAACCAATCTGTCAAAATACGAATTTTAAACCAACATCCATGTTCTT 180
Qy 61 LysLeuGlyValMetLeuValGlyTyrPGLyGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db 181 AATTTAGGGGTAAATGCTTGAGGTTGGGGTGGAAACAGGCTTCACCTCACCGGTGGT 240
Qy 81 ValIleAlaAsnArgGlyGlyIleSerTyrPalaThrLysAspLysIleGlnIleAsn 100
Db 241 GTATTATGCTAACCGAGAGGGCAATTTCATGGCTTACAAAGAGCAAGATTCACCAAGCAAT 360
Qy 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnIleGlyGlu 120
Db 301 TACTTTGGCTCCCTCCACCCCAAGCCTCAGCTATCCGAGTTGGTCTTCCAGGAGAGGAA 360
Qy 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 361 ATCTATAGCCCCCATTTCAAGACCTGCTTCCAAATGTTAACCTTACGACATGTGTTGGG 420
Qy 141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 421 GGATGGGATATCAGCAACATGAACCTGCTGATGCCATGGCCAGGCGAAAGCTTTGAC 480
Qy 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
Db 481 ATCGATTTGGCAGAAAGCAGTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAAATC 540
Qy 181 TyrAspProAspPheIleAlaAlaAsnGlnIleGluArgAlaAsnAsnValIleLysGly 200
Db 541 TATGACCCGATTTCAATGCTGTCACCAAGAGAGGTCACCAACACGTCATCAAGAGGC 600
Qy 201 ThrLysGlnIleGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnIleThr 220
Db 601 ACBAACCAAGACAGCAATTCACAAATCATCAAGACATCAAGCGCTTTAAGAAAGCACCC 660
Qy 221 LysValAspLysValValValLeuTyrPThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db 661 AAAGTGGACAAAGGTGGTGTGACTGTGAGACTGCCAACACAGAGAGTACATTAATTGGTT 720
Qy 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAlaAspArgAsnGluIleGln 260
Db 721 GTGGGCTTATATACACCATGAGAAATCTCTTGTGCTGTGACAGAAATGAGGCTGAG 780
Qy 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 781 ATTTCTCTTCCACCTTGTATGCCATTGCTGTGTTATGGAATGTCTCTTCATTATAT 840
Qy 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 841 GGAAAGCCCTCAGAAACCTTTGTACCAAGGCTGATTGATCTTCCATCGGAGAACACT 960
Qy 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 901 TTGATTGGTGGAGATACCTTCAGAGCTGTGACACCAAAATGAATCTGTGTGGTGTAT 960
Qy 321 PheLeuValGlyValGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db 961 TTCCTTGTGGGGCTGTATCAAGCCAAACATCTATATGCTCATCAACCACTTCGGGAAC 1020
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
Db 1021 AATGATGTATGATGATCTTTCGGCTCCCAAACTTTCGGTTCGAAGAAATCTCCAGAGG 1080
Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db 1081 AAGGTTGTGATATATGTGCAACACCAATGCAATCTCTATATAGCTGTGTGAACATCCA 1140
Qy 381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400

```

```

Db 1141 GACCATGTTGTTGTTATTAAGTATGCTTACAGGGGACAGCAAGAGCCATGAT 1200
Qy 401 GlnTyrThrSerGlnIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
Db 1201 GAGTACACTTCAGAGATATATTCATGGGTGGAAGAGCAACATGTTTGGACACACATGC 1260
Qy 421 GluAspSerLeuLeuAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1261 GAGATTTCCCTCTTACCTGCTCTTATTAATCTTGGACTTGGTCTTCTTGAGCTCAGC 1320
Qy 441 ThrArgIleGluPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
Db 1321 ACTAGATACGAGTTTAACTGTAATAATGAGGAAATTCACATTCATCCACCAAGTTGCT 1380
Qy 461 ThrIleLeuSerTyrIleThrLysAlaProLeuValProProGlyTyrProValValAsn 480
Db 1381 ACCATCTCAGCTTACCTCACCAGAGCTCTCTGTGTCCACCGGTACACACAGTGGAAAT 1440
Qy 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1441 GCATTTGCAAAGCAGGCTGCATGCTGGAACAAATATATAGGGCTGTGTGATTTGGCC 1500
Qy 501 ProGluAsnAsnMetIleLeuGlnTyrLys 510
Db 1501 CCAGAGATATACATGATTTCTCGAGTACAAAG 1530

RESULT 4
ADQ14490
ID ADQ14490 standard; cDNA; 1760 BP.
XX
AC ADQ14490;
XX
DE 23-SEP-2004 (first entry)
XX
XX wild type soybean myo-inositol 1-phosphate synthase cDNA #1.
XX
XX Soybean: myo-inositol 1-phosphate synthase; gene: ss;
XX myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate.
XX
XX Glycine max.
XX
XX
XX Key Location/Qualifiers
FH 54..1586
FT CDS /tag=a
FT /product="wild type soybean myo-inositol 1-phosphate
FT synthase #1"
PN US2004128713-A1.
XX
XX 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX 26-APR-1999; 99US-00299315.
XX 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX (GRAC/) GRACE D J.
XX (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX MPI; 2004-533135/51.
XX P-PSDB; ADQ14491.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.

```

XX Claim 4; SEQ ID NO 1; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
 CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
 CC phosphate. The invention also relates to a chimeric gene operably linked
 CC to suitable regulatory sequences, where expression of the chimeric gene
 CC results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase, or a soybean plant
 CC comprising the chimeric gene, a method of making the soybean plant, a
 CC seed of the soybean plant, a soy protein product derived from the
 CC processing of soybean seeds, a method of making or producing a soy
 CC protein product and a method of using a soybean plant homozygous for at
 CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 CC nucleic acid fragment and methods are useful for producing plants with
 CC decreased raffinose, stachyose and phytic acid content and increased
 CC sucrose and inorganic phosphate content, leading to valuable and useful
 CC soybean products. This sequence represents cDNA encoding a wild type
 CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.

SO Sequence 1760 BP; 494 A; 371 C; 411 G; 484 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,4e-253 Length: 1760
 Score: 2632.00 Matches: 510
 Percent Similarity: 100.00% Conservatave: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-10-718-952-2 (1-510) x ADQ14490 (1-1760)

QY 1 MetpneilegluasnphelysvalgiuCyeproasnvallystYrthrgiunhrgiulle 20
 DB 54 AGTTCATGAGAAATTTAAGGTGAGTGTCTTAAGTAGAATGACCCAGACTGAGATT 113
 QY 21 GlnSerValTYrAsnTYrGluThrThrgiuleuValHiegluasnargasnlyThrTYr 40
 DB 114 CAGTCGCTTACACTACGAAACCAACCACTTTTTCACGAGAACGAAATGGCACTTAT 173
 QY 41 GlnTrpIleVallyseBrolyserVallystYrGluPhelysThrasnIleHisValPro 60
 DB 174 CAGTGATGTCAAAACCAATCTGTCAATAGCAATTTAAACCAACATCCATGTTCT 233
 QY 61 LysLeuGlyValMetLeuValGlyTrpGlyAsnAsnGlySerThrLeuThrgly 80
 DB 234 AAATTAGGGGTAACTGTGGGTGGGTGGGAAACACGGCTCAACCTCCAGCGTGGT 293
 QY 81 ValIleAlaAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnIleAlaAsn 100
 DB 294 GTTATTGCTACCGAAGGGCATTTCAATGGCTCAAAAGCAAGATTCAACAGCCAT 353
 QY 101 TyrPheGlySerLeuThrgiulnAsnAlaIleArgValGlySerPheGlnGlyGlu 120
 DB 354 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCAGTTGGGCTTCCAGGAGAGAA 413
 QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
 DB 414 ACTTAAGCCCATTCAGAGCCGTCTTCAATGTGTAACTTGACCAACATGTGTGGG 473
 QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
 DB 474 GGATGGGATATTCGAAACATGAACCTGGCTGATGCCATGGCCAGGCAAAAGTGTGGC 533
 QY 161 IleAspLeuGlnLysGlnLeuArgProTYrMetGlySerMetLeuProLeuProGlyIle 180
 DB 534 ATCGATTTCAGAGAGAGTTGAGGCTTTCATGGAATTCATGTTCCACTCCCGGAATC 593
 QY 181 TYrAspProAspPheIleAlaAsnGlnGluGluArgAlaAsnAsnValIleLysGly 200
 DB 594 TATGACCCGAGATTTCATTGCTGTCCAAACAGAGAGCGTGCACAAACGTCAACAGGGC 653

QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
 DB 654 ACAAGCAAGAGCAAGTTCACAAATCATCAAAAGCATCAAGAGCGTTAAAGAACCAACC 713
 QY 221 LysValAspLysValValValLeuTrpPhraIaAsnThrgiulnArgTYrSerLysLeuVal 240
 DB 714 AAAGGAGCAAGGTGGTGTATGTGGACTGTGCACACAGAGAGGTACAGTAATTTGGTT 773
 QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsnGluAlaGlu 260
 DB 774 GTGGGCTTATGACACATGAGAAATCTTTGGCTGTCTGTGCACAGAAATAGAGCTGAG 833
 QY 261 IleSerProSerThrLeuTYrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
 DB 834 ATTCTCTCCATCACTTGTATGCCATTCGTTGTGTATGAAAATAGTTCTTTCAATTAT 893
 QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
 DB 894 GGAAACCCCTCAGACACCTTTGTACAGAGCGCTGATTGATCTTGCCATGCGGAGAACT 953
 QY 301 LeuIleGlyIAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
 DB 954 TTGATTGTGGAGATGACTTCAAGAGTGTCTCAGACCAAAATGAATCTGTGTGGTTGAT 1013
 QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTYrAsnHisLeuGlyAsn 340
 DB 1014 TTCCTTGGGGGGCTGTATCAAGCAACATGATATAGTCAGTTACCAACATCTGGGAAAC 1073
 QY 341 AsnAspGlyMetAsnLeuSerLAspProGlnThrPheArgSerLysGluIleSerLysSer 360
 DB 1074 AATGATGGTATGAATCTTGGCTCCACAACTTTCCTTCAAGAAATCTCCAAAGACC 1133
 QY 361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTYrGluProGlyGluHisPro 380
 DB 1134 AACGTTGTGATGATATGTGTCAACAGCAATGCCATCTTATGAGCCTGTGTAAATCCA 1193
 QY 381 AspHisValValIleLysTYrValProTYrValGlyAspSerLysArgAlaMetAsp 400
 DB 1194 GACCATGTTGTTATTATTAAGATATGCTTACGTAGGGGACAGACAGAGACCATGAT 1253
 QY 401 GlnTYrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
 DB 1254 GAGTACACTTCAGAGATATTCATGGGTGGAAAGACACCACTGTTTGGACAAACATGTC 1313
 QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
 DB 1314 GAGGATTCCTCTTATAGCTCTCTATATCTTGAGCTTGCTCTTGTCTGTAAGCTCAGC 1373
 QY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
 DB 1374 ACTAGATAGCAATTAAAGCTGAAATAGAGGAAATTCACATCTCCACCCAGTGTGCT 1433
 QY 461 ThrIleLeuSerTYrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
 DB 1434 ACCATCTTCACGTACTCTCAACAGGCTCCTGTGGTTCCACCGGATACCACTGTGTGAT 1493
 QY 481 AlaLeuSerLysGluArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
 DB 1494 GCATTGTCAAAACAGCGTGCATGCTGTGAAAACATATAGAGGCTTGTGTGGATTGGCC 1553
 QY 501 ProGluAsnAsnMetIleLeuGluTYrLys 510
 DB 1554 CCAGAGAAATACATGATTCGTGAGTACAAG 1583

RESULT 5

AAV62440 standard; cDNA; 1782 BP.

AAV62440;

17-OCT-2003 (revised)
 02-FEB-1999 (first entry)

XX Soybean wild-type myo-inositol 1-phosphate synthase cDNA.
 DE Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
 KW phytic acid; db.
 XX
 OS Glycine max, line LR13.
 XX
 PH Key Location/Qualifiers
 FT CDS 54..1586
 FT /*tag= a
 PN MO9845448-A1.
 XX
 XX 15-OCT-1998.
 PD
 XX 07-APR-1998; 98MO-US006822.
 PF
 XX 08-APR-1997; 97US-00835751.
 PR
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA
 XX Hitz WD, Sebaetian SA;
 PI
 XX WPI; 1998-568353/48.
 DR P-PSDB; AAW9740.
 XX
 PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
 PT for generating plants with altered levels of e.g. raffinose, stachyose,
 PT phytic acid, etc.
 XX
 PS Example 5; Page 44-45; 63pp; English.
 XX
 CC This is the nucleotide sequence of cDNA encoding the wild-type soybean
 CC myo-inositol 1-phosphate synthase (MI 1-PS) present in clone p5bm1-1ps
 CC (ATCC 97970). The clone was isolated from a cDNA library of soybean line
 CC LR13 by hybridisation to a probe made from MI 1-PS cDNA of Arabidopsis
 CC thaliana. MI 1-PS is involved in glucose metabolism to phytic acid,
 CC raffinose and stachyose. A mutant MI 1-PS nucleic acid (see AAW62443) has
 CC been identified in soybean line LR13, a mutagenised line of low raffinose
 CC saccharide phenotype. Sequencing revealed a single base change mutation
 CC (G to T at base 1241) in the LR13 sequence. The mutation results in a
 CC seed phenotype of very low raffinose saccharide sugars, very high sucrose
 CC and low phytic acid. The nucleic acid is used to alter the raffinose
 CC saccharide, sucrose, phytic acid and inorganic phosphate content of
 CC soybean seeds, leading to useful soybean products, e.g. a seed phytic
 CC acid content of less than 17 ug/g, a seed content of raffinose and
 CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
 CC greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)
 CC
 XX Sequence 1782 BP; 516 A; 371 C; 411 G; 484 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 2, 44e-253 Length: 1782
 Score: 2632.00 Matches: 510
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-10-718-952-2 (1-510) x AAW62440 (1-1782)
 QY 1 MetPheIIGluAsnPhelysValIGuCyPProAsnVallyrThrgluThrgluIle 20
 DB 54 ATGTTTCATCGAGAAATTTTAAGTGTGAGTGTCTTAAGTGAAGTGAACCGAGCTGAGATT 113
 QY 21 GlnSerVallyrAsnThrgluThrgluLeuValHISgluAsnArgAsnGlyThrTyx 40
 DB 114 CACTCGGTGACACTACGAAACACCGAACTTGTTCAAGAAACGAAATGACACTTAT 173
 QY 41 GlnTrpIleVallyrProlyrSerVallyrTyxGluPhelysThrsAnIleHisValPro 60
 DB 174 CAGTGATTTGTCAACCAATCTGTCAATATGAAATTTAAACCAATCCATGTTCTT 233

QY 61 LysLeuGlyValMetLeuValGlyTrpGlyGlyAsnAsnGlySerThrLeuThrglyGly 80
 DB 234 AAATTAGGGGTATATGTTTGTGGGTGGGTAACAAAGGCTCAACCTCCACGGGTGT 293
 QY 81 ValIleAlaAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
 DB 294 GTTATTGCTTACCGAAGGGCATTTTCAATGGCTACAAAGACAAAGTTCAACAGCCAT 353
 QY 101 TyrPheGlySerLeuThrglnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
 DB 354 TACTTGCTCCCTCCACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGGAGAGAA 413
 QY 121 IleTyrAlaProPheLysSerLeuLeuPProMetValAsnProAspAspIleValPheGly 140
 DB 414 ATCTATGCCCCATTCMAAGCCCTGCTTCAATGGTTAACTTCAGACAAATGTGTTGG 473
 QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
 DB 474 GGATGGGATATACGACAACTGAACCTGCTGATGCTATGCGCAAGGCAAGGTTTGAC 533
 QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
 DB 534 ATCATTTTCAGAAAGCAGTTGAGGCTTACATGGAATCCATGCTTCCATCCCGGAATC 593
 QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluArgAlaAsnAsnValIleLysGly 200
 DB 594 TATGACCCGGAATTTCACTTGCTCCCAACGAAGAGCGTGCCAAACCACTCATCAAGGGC 653
 QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysValPheLysGlnAlaThr 220
 DB 654 ACMAAGCAAGACCAAGTTCACAAATATCAAAAGACATCAAGGCTTTAAGAAAGCAC 713
 QY 221 LysValAspLysValValValLeuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
 DB 714 AAAGTGACAAAGTGGTGTATCTGTCGTCGCAACACAGAGAGTACGTAATTTGGTT 773
 QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaIleValAspArgAsnGluValGlu 260
 DB 774 GGGGGCTTAAAGACCACTGAGAAATCTTGGCTGCTGTGACGAATGAGGCTGAG 833
 QY 261 IleSerProSerThrLeuTyrAlaIleAlaCyValMetGluAsnValProPheIleAsn 280
 DB 834 ATTTCTCTTCCACTTGTATGCCATTGCTTGTATGAAATGTTCTTTCAATTAAT 893
 QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
 DB 894 GGAAGCCCTCAGAACACTTTGTACAGAGGCTGATTGATCTTGCCATGCGAGAACACT 953
 QY 301 LeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
 DB 954 TTGATTGGTGGAGATGACTTCAAGAGTGTACAGACCAAAATGAATCTGTGTTGAT 1013
 QY 321 PheLeuValGlyAlaGlyIleLysPProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
 DB 1014 TTCTCTTGGGGCTGGTATCAAGCCACATCTTATGTAGTTCAACCACTTGGGAAAC 1073
 QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyIleSerLysSer 360
 DB 1074 AATGATGGTATGATCTTCCGCTCCACAAACTTTCCTCCAAAGGAATCTCCAAAGAC 1133
 QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
 DB 1134 AACGTTGTGATGAATATGTGTCAACGCAATGCACTTCTATGAGCTGGTGAACATCA 1193
 QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
 DB 1194 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAAGGGAACGAAAGAGCCATGAT 1253
 QY 401 GlnTyrThrSerGluIlePheMetGlyLysSerThrIleValLeuHisAsnThrCyx 420
 DB 1254 GAGTACACTTCAGAGATATTCATGGGTGGAAGACACACATGTTTTCGCAACACATGC 1313


```
QY 421 GIuAspSerLeuLeuAlaIaIaProIleIleLeuAspLeuValIleuLeuAlaIuLeuSer 440
DB 1314 GAGGATTCCTCTTACTCTGCTTATTAATCTTGGAATCTTGCTTCTTGCTGAGCTCAAC 1373
QY 441 ThrArgIleGluPheLysAlaGluAenGluGlyLysPheHisSerPheHisProValAla 460
DB 1374 ACTAGATATGAGATTAAAGCTGAAATGAGGGGAAATTCACACATTCACACCGAGTTGCT 1433
QY 461 ThrIleLeuSerTyrLeuThrLysAlaIaProLeuValProProGlyThrProValValAsn 480
DB 1434 ACCATCTCAAGCTACTCTACCAAGGCTCTCTGCTTCCACCGGATACACCAAGTGAT 1493
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1494 GCATTGTCAAGCAGCGCTCAATGCGAAACATATGAGGGCTTGTTGATTGGCC 1553
QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
DB 1554 CCAGAAATACATGATCTCGAGTCAAG 1583

RESULT 6
AAV62443
ID AAV62443 standard; cDNA; 1533 BP.
XX
AC AAV62443;
XX
DT 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
XX
DE Soybean mutant myo-inositol 1-phosphate synthase cDNA.
XX
KM Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KW phytic acid; de.
XX
OS Glycine max; line LR33.
XX
PN WO9845448-A1.
XX
PD 15-OCT-1998.
XX
PF 07-APR-1998; 98WO-US006822.
XX
PR 08-APR-1997; 97US-00835751.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E. I.
PI Hitz WD, Sebastian SA;
XX
XX WPI; 1998-568353/48.
DR P-PSDB; AAW79741.
XX
PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
PS phytic acid, etc.
XX
XX Example 5; Page 48-49; 63p; English.
XX
CC This is the nucleotide sequence of cDNA encoding a mutant soybean myo-
CC inositol 1-phosphate synthase (MI 1-PS) present in clone LR33-10. This
CC clone was isolated from a cDNA library of soybean line LR33 by PCR
CC amplification (see AAV62441-42). Line LR33 was obtained by chemical
CC mutagenesis of wild-type soybean genome and as a reduced raffinose
CC saccharide phenotype. MI 1-PS is involved in glucose metabolism to phytic
CC acid, raffinose and stachyose. Sequencing revealed a single base change
CC mutation (G to T at base 1241) in the LR33 sequence when compared to the
CC wild-type sequence (see AAV62440). The mutation results in a seed
CC phenotype of very low raffinose saccharide sugars, very high sucrose and
CC low phytic acid. The mutated MI 1-PS nucleic acid is used to alter the
CC raffinose saccharide, sucrose, phytic acid and inorganic phosphate
CC content of soybean seeds, leading to useful soybean products, e.g. a seed
CC phytic acid content of less than 17 ug/g, a seed content of raffinose and
CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
CC greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)
```

```
XX
SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6,266-253
Score: 2627.00
Percent Similarity: 99.80%
Best Local Similarity: 99.80%
Query Match: 99.81%
DB: 2 Gaps: 0

US-10-718-952-2 (1-510) x AAV62443 (1-1533)
QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 1 ATGTTGATGAGAAATTTTAAGGTGAGGTGCTCTAATGGAATGATACCCAGACTGAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCCGTGTACACTACGAAACCAACCGAACTGTTCACGAGAACGGAATGGACCTAT 120
QY 41 GlnTrpIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
DB 121 CAGTGGATTGTCAAAACCAAAATCTGTCAATATCGAATTTAAACCAACATTCATGTTCT 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyLysAsnAsnGlySerThrLeuThrGlyGly 80
DB 181 AAATTAGGGGTAAATGCTTGTTGGGTGGGTGAAACAAACGGCTCAACCTCAACCGTGGT 240
QY 81 ValIleAlaAsnArgGluGlyLysSerTrpAlaThrLysAspLysIleGlnIleAlaAsn 100
DB 241 GTTATTGCTTAACCGAGAGGCAATTCATAGCGCTACAAAGCAAAATTCACAAACGCAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTTGGCTCCCTCAACCCAGCTTAGCTATCCAGTTGGGTCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCATTCACAGAGCGCTGCTCCAAATGTTAACCCTGACGACATGTGTTGGG 420
QY 141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGATATCAGCAACATGAACTGCTGATGTCATGCGCAGGCAAAAGGTGTTGAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGlnSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTGCAAGAGCATTTGAGGCTTACATGAAATTCATGCTTCCACTCCCGGAAATC 540
QY 181 TyrAspProAspPheIleAlaIaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 541 TATGACCGGATTTTCATTTGCTGCCAAACAAAGAGGCGGTGCAACACGTATCAAGGGC 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 601 ACAAAGCAAGACCAAGTTCACAAATCATCAAAAGCATCAAGCGCTTTAAGAAAGCACAC 660
QY 221 LysValAspLysValValValLeuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGGAAGAGGTGTTGTACTGTGAGCTGCCAACAGAGAGTACGTAAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu 260
DB 721 GTGGGCTTTAAGACCACTGAGATCTCTTGCGTGGCTGTGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCTCTTCCACCTTGATCCATTTGCTGTGTATGAAAAGTTCTTTCAATTAAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAAGAACATTTGTACAGGGCTGATTTGATCTTGCCATCGCGAGAACT 900
```



```

QY 301 LeuileG1yAspAspPheylsSerGlyGlnThrlySerValleuValAsp 320
DB 901 TTGATGCTGGAGATGACTTCAAGAGTGTGACCAAAATGAAATCTGTTGGTTGAT 960
QY 321 PheLeuValGlyValaglylylelyserProthSerileValserTyranhiLeuGlyAsn 340
DB 961 TTCCTTGGGGGCTGGTATCAAGCCAACTATAGTCAGTTAACAACATCTGGGAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerlyseGlylyleSerlySer 360
DB 1021 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaileuTyrglyProGlyGlylyleSerPro 380
DB 1081 AACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 381 AspHisValValValilelyTyrglyProTyrglyValGlyAspSerlyAspAlaMetAsp 400
DB 1141 GACCATGTTGTTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 401 GlyTyrglyThrSerGlylylePheMetGlylylyleSerThrileValleuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAGATATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 421 GluAspSerleuValleuAlaProilleleuAspPheValleuValleuValleuSer 440
DB 1261 GAGGATTTCCCTTACGCTGCTCTATATTCATGATGATGATGATGATGATGATGATGAT 1320
QY 441 ThrArgileGlyPheValAspGlylyleGlylylylePheHisSerPheHisProValAla 460
DB 1321 ACTAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 461 ThrileuSerTyrglyleuThrlyValProleuValProProGlylyleProValValAsn 480
DB 1381 ACCATCTCAGCTACCTCAACCAAGGCTCTCTGCTTCCACCGGCTACCAAGTGGTGAAT 1440
QY 481 AlaLeuSerlyGlnArgAlaMetleuGluAsnileMetArgAlaCysValGlyleuAla 500
DB 1441 GCATTTCAAGAGGCTGCAATGCTGCAAAACATTAATGAGGCTGTGATGGCC 1500
QY 501 ProGluAsnAsnMetileuGlylyleGlylylyle 510
DB 1501 CCAGAGATATACATGATTTCTCGATGATCAAG 1530
RESULT 7
ADQ14494
ID ADQ14494 standard; cDNA; 1533 BP.
XX
AC ADQ14494;
XX
XX 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #1.
XX
KM Soybean: myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KM inorganic phosphate; mutant.
XX
OS Glycine max.
XX
XX Synthetic.
XX
FH Key 1.1533 Location/Qualifiers
FT CDS 1.1533
FT /tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT /synthase #1"
FT mutation replace(1241,G)
FT /tag= b
PN US2004128713-A1.
XX
XX 01-JUL-2004.
XX

```

```

PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-0029315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEEA/) SEEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STREIT/) STREIT L G.
XX
PI Hiltz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX WPI; 2004-533135/51.
XX P-PDB; ADQ14495.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
PS Claim 10; SEQ ID NO 5; 48bp; English.
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 6,26e-253 Length: 1533
Score: 2627.00 Matches: 509
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.81% Indels: 0
DB: 12 Gaps: 0
US-10-718-952-2 (1-510) x ADQ14494 (1-1533)
QY 1 MetPheileGluAsnPheylsValGlyCysProAsnVallyTyrglyThrGlylyle 20
DB 1 ATGTCATCGAGAAATTTTAAAGTTGATGCTCTTAATGTCAAGACCAACGACTGAT 60
QY 21 GlnSerValTyrglyAsnTyrglyThrGlylyleValleuValleuValleuValleuVal 40
DB 61 CAGTCGCTGATCACTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCACTAT 120
QY 41 GlnTrileValleuValleuValleuValleuValleuValleuValleuValleuValleu 60
DB 121 CAGTGAATGTCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCACTAT 180
QY 61 LysLeuGlyValMetLeuValGlylyleTyrglylyleValleuValleuValleuValleu 80
DB 181 AATTTAGGGGATATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY ValileleAsnArgGlylylyleSerTyrglyThrlyAspPheylleGlnGlnAlaAsn 100

```

Db	241	CTTATTCGTAAACCGAGAGGGCATTTTCATGGGCTACAAAGACAAAGATTTCACAAAGCCAAAT	300
Qy	101	TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu	120
Db	301	TACTTTGGCTCCCTTCACCCAGAGCCCTACAGCTATCCGAGTTGGGCTCTTCCAGGGAGAGAA	360
Qy	121	IleTyrAlaProPheIysSerLeuLeuProMetValAsnProAspAspIleValPheGly	140
Db	361	ATCTATGCCCATTCACAGAGCCCTGCTCCCAATGGTTAAACCTTGACGACATGTGTTGGG	420
Qy	141	GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIysValPheAsp	160
Db	421	GGATGGGATACAGACACATGAACTGGCTATGCTCATGGCCAGGCGAAAGGTGTTGGAC	480
Qy	161	IleAspLeuGlnGlyGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyIle	180
Db	481	ATCGATTTGCAGAAAGCAGTTGAGGGCTTACATGGATCATGCTTCCATCCCGGAAATC	540
Qy	181	TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleIysGly	200
Db	541	TATGACCCGGATTTTCATTGCTGTGCCAACCAAGAGGCGCTGCCAACACGTCATCAAGGCG	600
Qy	201	ThrIysGlnGlnGlnValGlnGlnIleIleIysAspIleIysAlaPheIysGlnAlaThr	220
Db	601	ACAAAGCAGAGCAGATTCACAAATATCTCAAGACATCAAGGCGTTTAAGAGAGCCACC	660
Qy	221	LysValaAspLysValaValIleuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal	240
Db	661	AAAGTGACCAAGTGCGTGTGTACTGTGACCTGCCAACAGAGAGGTACAGTAATTTGGTT	720
Qy	241	ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValaAspArgAsnGluAlaGlu	260
Db	721	GTTGGCCCTTAATGACACATGAGAGATCTCTGGCTGTGGACAGAAAGAGGCTGAG	780
Qy	261	IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn	280
Db	781	ATTTCCTCTTCCACCTGTATGCCATTCCTGTGTATGAGAAATGTTCTCTTCAATTAT	840
Qy	281	GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr	300
Db	841	GGAGGCCCTCAGAACACTTTTGTACCAAGGCGTGAATGATCTTGGCATGCGAGGAACCT	900
Qy	301	LeuIleGlyGlyAspAspPheIysSerGlyGlnThrIlybMetLysSerValLeuValAsp	320
Db	901	TTGATTTGGTGAGATGACCTTCAGAGTGTGTGACACCAAAAGAAATCTGTGTGGTTGAT	960
Qy	321	PheLeuValGlyValaGlyIleIysProThrSerIleValSerTyrAsnIleLeuGlyAsn	340
Db	961	TTCTCTTGGGGGCGGTGATCAAGCAACATCTATAGCTATACAAACATCTGGGAAC	1020
Qy	341	AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIysGluIleSerLysSer	360
Db	1021	AATGATGGTATGAATCTTTCGGCTCCCAAACTTCCGTTCCAGAGAAATCTCCAAAGAC	1080
Qy	361	AsnValValaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro	380
Db	1081	AACGTTGTTGATGATATGTCACACGACATGCCATCCTCTATGAGCCGTGGAAATCTCA	1140
Qy	381	AspHisValValaValIleIleTyrTyrAlaProTyrTyrAlaGlyAspSerLysArgAlaMetAsp	400
Db	1141	GACCAATGTTGTTGTTAATTAAGATAGTGCCTTACGTAAGGAGAACAGCAATGAGCCATGAT	1200
Qy	401	GluTyrThrSerGluIlePheMetGlyIlyLysSerThrIleValLeuHisAsnThrCys	420
Db	1201	GAGTACACTTCAGAGATATTCATGGGTGGAAAGACACATTTGTTTGCACACACATGC	1260
Qy	421	GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer	440
Db	1261	GAGGATTCCTCTTATGCTGTCTCTATATCTTGACCTTGCTCTTCTTGACACTGAC	1320
Qy	441	ThrArgIleGluPheIlyAlaGluAsnGluIlyLysPheHisSerPheHisProValAla	460
Db	1321	ACTGAAATCTGAGTTTAAAGCTGAAATATAGGGAATAATTCATCTACCTCCACCGATGGCT	1380

Oy	461	ThrIleuSerTYrLeuThrIyValAlaProLeuValProProGlyThrProValIValAsn	480
Db	1381	ACCATTCTTCAGCTAACCTCCACCAAGGCTCTCTGCTTCACCGGGTACACAGGTGTAAAT	1440
Oy	481	AlaIeuSerIyGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla	500
Db	1441	GCATTGTCMAAGCAGCGCTGCATTCCTGGAAAACATATATGAGGCTGTGTGATTGGCC	1500
Oy	501	ProGluAsnAsnMetIleLeuGluTYrLys	510
Db	1501	CCAGAGATTAACATGATTCTCGATTACAAG	1530
RESULT 8			
ADS81997	ID	ADS81997 standard; cDNA; 1533 BP.	
XX	AC	ADS81997;	
XX	DT	18-NOV-2004 (first entry)	
DE		Soybean myo-inositol 1-phosphate synthase cDNA mutant #1.	
XX		Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;	
KW		raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.	
XX		Glycine max; line LR33.	
OS		Synthetic.	
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..1533	
FT		/*tag= a	
FT		/product= "myo-inositol 1-phosphate synthase"	
FT	mutation	replace(1188,G)	
FT		/*tag= b	
XX			
PN		US2003074685-A1.	
PD		17-APR-2003.	
XX			
PF		11-MAR-2002; 2002US-00025003.	
PR		08-APR-1997; 97US-00835751.	
PR		07-APR-1998; 98WO-US006822.	
XX			
PA		(HITZ/) HITZ W D.	
PA		(SEBA/) SEBASTIAN S A.	
PI		Hitz WD, Sebastian SA;	
XX			
DR		WPI; 2004-639957/62.	
DR		P-P5DB; ADS81998.	
XX			
PT		Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-	
PT		phosphate synthase, useful for altering raffinose saccharide, sucrose,	
PT		phytic acid and inorganic phosphate content of soybean seeds.	
XX			
PS		Claim 8; SEQ ID NO 5; 34pp; English.	
XX			
CC		The invention relates to an isolated nucleic acid fragment encoding a	
CC		soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-	
CC		phosphate synthase having decreasing capacity for the synthetas for myo-	
CC		inositol-1-phosphate. Also included are a chimeric gene (comprising the	
CC		nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its	
CC		complement, subfragment or the complement of the subfragment, operably	
CC		linked to suitable regulatory sequences, where expression of the chimeric	
CC		gene results in a decrease in expression of an endogenous or native gene	
CC		encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant	
CC		comprising the chimeric gene (with a heritable phenotype of a seed phytic	
CC		acid content of less than 17 micromol/g, a seed content of raffinose plus	
CC		stachyose of less than 14.5 micromol/g, and a seed sucrose content of	
CC		greater than 200 micromol/g, provided that the plant is not LR33), seeds	
CC		from the plant, making a soybean plant with the heritable phenotype	

CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homozygous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with a
 CC heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence encodes a mutant myo-
 CC inositol 1-phosphate synthase.

XX Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6,266-253	Length:	1533
Score:	2627.00	Matches:	509
Percent Similarity:	99.80%	Conservative:	0
Best Local Similarity:	99.80%	Mismatches:	1
Query Match:	13	Indels:	0
DB:	13	Gaps:	0

US-10-718-952-2 (1-510) x ADS81997 (1-1533)

QY 1 MetPhe11eGluAsnPhelYsVal1GluCyPProAsnVal1YsThrGluThrGlu11e 20
 DB 1 AGTTTATGAGAAATTTTAAGTTGAGTGTCTTAATGTGAATGACCGAGATGAGATT 60
 QY 21 GlnSerVal1YsAsnTYrGluThrGluLeuVal1HieGluAsnArgAsnGlyThrTYr 40
 DB 61 CAGTCCGCTACAACTACGAAACACCACTGTTCAAGAAACAGAAATGCGACTAT 120
 QY 41 GlnThr11eVal1YsPro1YsSerVal1YsTYrGluPhelYsThrAsn11e1sValPro 60
 DB 121 CAGTGGATGTCAAACCCAAATGTCAAATACGAAATTTAAACCAACATCCAGTTCT 180
 QY 61 LysLeuGlu1YsMetLeuVal1GlyTYrGly1YsAsnAsnGlySerThrLeuThrGly1Y 80
 DB 181 AAATTAGGGGTAATGTTGTGGGTGGGTAACAACCGCTCAACCTCAGCGGTGCT 240
 QY 81 Val11eAlaAsnArgGluGly1YsSerTYrAlaThrLysAsp1Ys11eGlnGlnAlaAsn 100
 DB 241 GTTATTGCTAACGAGAGGCAATTCATGGGCTAACAAAGCAAGATTCACAAAGCCAA 300
 QY 101 TYrPheGlySerLeuThrGlnAlaSerAla11eArgVal1GlySerPheGlnGlyGlu 120
 DB 301 TACTTGGCTCCCTCCACCAAGCTCAGCTATCCGAGTTGGGCTTCCAGGAGAGGAA 360
 QY 121 11eTYrAlaProPhe1YsSerLeuLeuProMetVal1AsnProAspAsp11eVal1PheGly 140
 DB 361 ATCTATGCCCATCTCAAGAGCCCTGTTCCAATGTTAAACCTGACACATGTGTTGGG 420
 QY 141 GlyTYrAsp11eSerAsnMetAsnLeuAlaAspAlaMetAlaArgAla1YsVal1PheAsp 160
 DB 421 GGATGGGATATCAGCAATGAACTGGCTGATGCAATGGCCAGGCAAAAGGTTTGAC 480
 QY 161 11eAspLeuGln1YsGlnLeuArgProTYrMetGluSerMetLeuProLeuProGly11e 180
 DB 481 ATCGATTTCAGAAAGAGTTGAGGCTTACATGAAATCCATGTTCCCTCCCGGAATC 540
 QY 181 TYrAspProAspPhe11eAla1AsnGlnGluGluArgAlaAsnAsnVal11e1YsGly 200
 DB 541 TATGACCCGAGATTTCATCTGCTGCCAACAGAGAGACGTCACAACTCATCAAGGCG 600
 QY 201 ThrLysGlnGluGlnVal1GlnGln11e11eLysAsp11eLysAlaPhe1YsGluAlaThr 220
 DB 601 ACAAAAGCAAGACAGTTCAACAAATCATCAAAAGATCAAGAGCGTTTAAAGAAAGCCACC 660

QY 221 LysValAspLysVal1Val1LeuTYrThrAlaAsnThrGluArgTYrSerAsnLeuVal 240
 DB 661 AAAGTGACAAAGGTGGTGTACTGTGAGCTGCCAACACAGAGAGTACAGTAAATTTGGTT 720
 QY 241 Val1GlyLeuAsnAspThrMetGluAsnLeuVal1Ala1ValAspArgAsnGlu1YsGlu 260
 DB 721 GTGGGCTTAATGACACCATGGAGAAATCTTGCGCTGCTGTGGACAGAAATAGCGCTGAG 780
 QY 261 11eSerProSerThrLeuTYrAla11eAla1YsValMetGluAsnVal1ProPhe11eAsn 280
 DB 781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGTATGAAATATGTTCTTTCATTAAAT 840
 QY 281 GlySerProGluAsnThrPheVal1ProGlyLeu11eAspLeuAla11eAlaArgAsnThr 300
 DB 841 GGAAGCCCTCAGAAACCTTTGTACAGGAGCTGATTTGATTCCTTCATCGAGAGAACT 900
 QY 301 Leu11eGly1YsAspAspPhe1YsSerGlyGlnThrLysMetLysSerVal1LeuValAsp 320
 DB 901 TTGATTGGTGAAGATGACTTCAAGAGTGTCCAGACCAAAATGAAATCTGTGTTGAT 960
 QY 321 PheLeuVal1GlyAlaGly11eLysProThrSer11eVal1SerTYrAsnH1sLeuGlyAsn 340
 DB 961 TTCTTGTGGGGCTGGTATCAAGCCCAATCTATAGTCAACCAATCTGGGAAAC 1020
 QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlu11eSerLysSer 360
 DB 1021 AATGATGATGATGAATCTTGGCTCCAAACATTTCTCGTTCCAAAGAAATCTCCAGAGC 1080
 QY 361 AsnVal1YsAspAspMetVal1AsnSerAsnAla11eLeuTYrGluProGlyGluH1sPro 380
 DB 1081 AAGCTTGTGATGATATGATGTCACACAGCATGCAATCTCTTAAGAGCTGTGAACATCCA 1140
 QY 381 AspH1sVal1Val11eLysTYrVal1ProTYrVal1GlyAspSer1YsArgAlaMetAsp 400
 DB 1141 GACCATGTTGTTTATTAATGATATGCTTACATGAGGAGCAAGCAATGAGCCATGAT 1200
 QY 401 GlyTYrThrSerGln11ePheMetGly1YsSerThr11eVal1LeuH1sAsnThrCys 420
 DB 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGACACCATGTTTGCAACACATGC 1260
 QY 421 GluAspSerLeuLeuAla1YsPro11e11eLeuAspLeuVal1LeuLeu1YsSer 440
 DB 1261 GAGGATCCCTCTTACGCTCTCTATTAATCTTGGAATGCTGCTTCTTCTAGCTCAAC 1320
 QY 441 ThrArg11eGluPhe1YsAlaGluAsnGly1YsPheH1sSerPheH1sProVal1Ala 460
 DB 1321 ACTAGATGAGATTAAAGCTGAAATGAGGAAATTCACCTCATTTCCACCAGTTGCT 1380
 QY 461 Thr11eLeuSerTYrLeuThr1YsAlaProLeuVal1ProProGly1YsThrProVal1YsAsn 480
 DB 1381 ACCATCTCAGCTACCTCACCAAGGCTCTGTGTTCCACCGAGTACACCAAGTGTGAAT 1440
 QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsn11eMetArgAla1YsVal1GlyLeuAla 500
 DB 1441 GCATTGTCAAAACACAGCTGCATGCTGGAACATTAATAGAGGCTTGTGTGATTGGCC 1500
 QY 501 ProGluAsnAsnMet11eLeuGluTYrLys 510
 DB 1501 CCAGAGAAATTAATGATATCTCGAGTACAAAG 1530

RESULT 9
 ADQ14504
 ID ADQ14504 standard; cDNA; 1533 BP.
 XX ADQ14504;
 AC
 XX
 XX
 DT 23-SBP-2004 (first entry)
 XX
 DE Wild type soybean myo-inositol 1-phosphate synthase cDNA #2.
 XX
 KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
 KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
 KM inorganic phosphate.

XX Glycine max.
OS
XX
PH Key Location/Qualifiers
FT CDS 1..1533
FT /*tag= a
FT /product= "Wild type soybean myo-inositol 1-phosphate
FT synthase #2"
XX
XX
XX US2004128713-A1.
XX
XX 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX 26-APR-1999; 99US-00299315.
XX 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX (GRAC/) GRACE D J.
XX (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX MPI; 2004-533135/51.
XX P-PSDB; ADQ14505.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX Claim 4; SEQ ID NO 15; 48pp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents cDNA encoding a wild type
XX soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
XX SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 3,17e-250 Length: 1533
XX Score: 2600.00 Matches: 503
XX Percent Similarity: 99.41% Conservative: 4
XX Best Local Similarity: 98.61% Mismatches: 3
XX Query Match: 98.78% Indels: 0
XX DB: 12 Gaps: 0
XX
XX US-10-718-952-2 (1-510) x ADQ14504 (1-1533)
XX
XX QY 1 MetPheIleGluAsnPhelLysValGluCysProAsnValIleuThrGluThrGluIle 20
XX Db 1 ATGTTATCGAGAAATTTTAAAGGAGAGCTCTTAAGTGAAGTCAACCGAGACTGAGATT 60
XX QY 21 GlnSerValIleuAsnIleuThrGluThrGluLeuValIleGluAsnArgAsnGlyThrTyr 40

Db 61 CAGTCGCTGTAACAATAAGAAACCAACGAACTTGTTCAGAGAACAGAAATGGCACTAT 120
QY 41 GlnTrpIleValIleProLysSerValIleuThrGluPheLysThrAsnIleGhiValPro 60
Db 121 CAGTGGATGTCAGAAACCAAAATCGTCAATCACTAATTTAAACCAACCACTATTTCCA 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyLysAsnAsnGlySerThrLeuThrGlyGly 80
Db 181 AAATTGGGGGGGATGCTGTTGGGGTTGGGGTGAACCAACGGCTTACCTCAACCGGTGAT 240
QY 81 ValIleAlaAsnArgGluGlyTyrIleSerTrpAlaThrLysAspLysIleGlnIleAlaAsn 100
Db 241 GTTATTCCTAACAGAGAGGCAATTCATGAGGTTCMAAGAGCAACAAATCAACCAACCAAT 300
QY 101 TyrPheGlySerLeuThrGluAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 301 TACTTTGGCTCCCTTACCAACCAAGCTCAAGCTTTCAGATTTGAGTCTTCCAGAGAGAGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 361 ATCTATGCCCATTCAGAGAGTCTGCTTCCAAATGTTAATCCTGACGACATTTGTTGGG 420
QY 141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 421 GGATGGAGTATCAGCAACATGAACTGGCTGATGCAATGCGCAAGGCAAAAGGTGTTGAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
Db 481 ATCGATTGCGAAGACAGATGAGGCTTTCATGAAATCAATGTTCCACTCCCGCAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
Db 541 TACGACCCGGATTTTCAATGCTGCTGCCAACCAAGAGAGCGCGCAACACGTGATTAGGGC 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
Db 601 ACAAAGCAAGAGCAAGTTCAGCAAAATCATCAAAAGCATCAAGGCTTTAAAGCAACCAAC 660
QY 221 LysValAspLysValValValLeuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db 661 AAAGTGAACAGAGTGTGTCTGTGGACTGCCAACAGAGAGTATAGCAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluGlu 260
Db 721 GTAGGCTTAATGACCAATGAGATCTCTTGCTGCTGTGAGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaLysValMetGluAsnValProPheIleAsn 280
Db 781 ATTTCTCTCCACCTTGATATGCCATTGCTGTGATGGAAGAAATGTTCTTTCAATTAAAT 840
QY 281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 841 GGAAGCCCTCAACAACCTTTTATCAAGGCTGATTTGATCTTGCCATCGCGAGAACT 900
QY 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 901 TTGATTTGGTGGAGATGACCTTCAAGAGTGTGACGACCAAAATGTAATCTGTGTTGAT 960
QY 321 PheLeuValGlyAlaGlyTyrLysProThrSerIleValSerTyrAsnHileuGluLys 340
Db 961 TTTCTTGGGGGGCTGGTATCAAGCAACATCTAATAGTTAATCAACCATCTGGGAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db 1021 AATGATGGTAAATCTCGGCTCCAAACCTTCGCTCCAAAGAAATCTCCAAAGACC 1080
QY 361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db 1081 AACGTTGTGAGAGATGCTCAACAGCAATGCCATCTTAAGACCTGGTGAACATGCC 1140
QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetLysP 400

541 TACGACCCGGATTTCATTCCTGCCAACCAAGAGGCGTCCACAAACGATTAAGGCG 600
201 ThrLysGlnGluGluValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
601 ACAAGACAGACAGAGCTTCCAGCAATCATCAAGACATCAAGAGCGTTTAAAGAGCCACC 660
221 LysValAspLysValValValLeuTrpThrAlaAsnThrGluArgLysSerAsnLeuVal 240
661 AAAAGGGAACAAGGTGGTTCCTCGTGGACTGGCAACAGAGAGGATATGCAATTTGGTT 720
241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsnGluAlaGlu 260
721 GTAGGCGCTTAATGACCACTGAGAAATCTCTGGCTGCTGTGACAGAAATAGGCGTGA 780
261 ILeSerProSerThrLeuTrpAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
781 ATTTCTCTTCCACCTTGATGATGCCATTCCTGTGATGAAAGAGTTCTTTCATTAAT 840
281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
841 GGAAGCGCCCTCAGAACACTTTTGGACAGGCGTGAATGATCTTGCCATCGCAGAGAACT 900
301 LeuIleGlyGlyAspAspPheLysSerGlyGlnTrpLysMetLysSerValLeuValAsp 320
901 TTGATTGGTGAGATGACTTCAAGAGTGGTCAGACCAAAATGAAATCTGTGTGGTGGAT 960
321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
961 TTTCTTGTGGGGGCTGGTATCAAGCCACATCTATAGTTACACACCATCTGGGAAC 1020
341 AsnAspGlyMetAsnLeuSerAlaProGlnTrpPheArgSerLysGluIleSerLysSer 360
1021 AATGATGATGATGATCTCGGCTCCACAAACCTTCCTCCCAAGGAATCTCCAAAGAGC 1080
361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTTATGAGCGTGTGAACATCC 1140
381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
1141 GACCATGTTGTTTATTAAGTATGTGCTTACGTAGGGGATAGCAGAGAGCCATGAT 1200
401 GlyTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
1201 GAGTACACTTCAAGATATTCATGGGTGGAAGAACACCATTTGTTGCACAAACATGT 1260
421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
1261 GAGGATTCCTCTTTAGCTGCTCTATTAATCTTGGACTTGCTCTTGTCTAGAGCTGAGC 1320
441 ThrArgIleGluPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
1321 ACTAAGATCAAGTTTAAAGTAAATGAGGAAATTCACATCACTCCACCATGTTGCT 1380
461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
1381 ACCATCTCAGCATATCTGACCAAGGCTCTCTGTTGCCACCGGTTACACCATGTGTAAT 1440
481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaLysValGlyLeuAla 500
1441 GCAATGTCAAGAGGCTCAATGCTGGAACCATATAGAGGCTGTGTGTGATTTGGGCC 1500
501 ProGluAsnAspMetIleLeuGlyTyrLys 510
1501 CCAAGAAATTAACATGATTTCTCGAGTACAAAG 1530

RESULT 11

ADS82005 ID ADS82005 standard; cDNA; 1533 BP.

AC ADS82005;

XX 18-NOV-2004 (first entry)

XX DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 2.
XX KW Soybean; plant; myo-inositol 1-phosphate synthase; ssr gene; phytic acid;
XX KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX OS Glycine max; cultivar Wye.
XX FT Key Location/Qualifiers
XX FT CDS 1..1533
XX FT /*tag= a
XX FT /product= "myo-inositol 1-phosphate synthase"
XX PN US2003074685-A1.
XX PD 17-APR-2003.
XX PF 11-MAR-2002; 2002US-00025003.
XX PR 08-APR-1997; 97US-00835751.
XX PR 07-APR-1998; 98WO-US006822.
XX PA (HITZ/) HITZ W D.
XX PA (SEBA/) SEBASTIAN S A.
XX PI Hitz WD, Sebastian SA;
XX DR MPI: 2004-639957/62.
XX DR P-PsDB; ADS82006.
XX PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
XX PS Claim 2; SEQ ID NO 15; 34pp; English.
XX CC The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of native gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homologous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.
XX SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Alignment Scores: 3 176-250 Length: 1533
Pred. No.: 2600.00 Matches: 503
Score: 99.41% Conservative: 4
Percent Similarity:

Best Local Similarity: 98.63% Mismatches: 3
 Query Match: 98.78% Indels: 0
 DB: 13 Gaps: 0
 US-10-718-952-2 (1-510) x ADS82005 (1-1533)

QY 1 MetPhe1EGluAspPheLeuValGluCysProAsnValIleuThrGluThrGlu1 20
 DB 1 AGTTTCATGAGAAATTTTAAAGTAGAGAGCTCTAAATGTGAAGTACCGAGAGCTGAGATT 60
 QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHISGluAsnArgAsnGlyThrTyr 40
 DB 61 CAGTCGCTGACAACTACGAAACACCGAAGCTTTCACGAGAAACAGAAATGCACTTAT 120
 QY 41 GlnTP1LeuValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
 DB 121 CAGTGATGTCAAAACCAATCCCTCACTACCAATTTAAACCAACCCCACTGTTCGA 180
 QY 61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGly 80
 DB 181 AAATGGGGGTGATGCTGTGGTGGGTGGGAAACAGGCTCTACCTCACCGGTGGT 240
 QY 81 ValIleAlaAsnArgGluGlyIleSerTyrAlaThrLysAspLysIleGlnIleAsn 100
 DB 241 GTTATTGCTAACAGAGAGGCAATTCATGGGCTACAAAGGACAAAGATTCAACAGCCAT 300
 QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
 DB 301 TACCTTGGCTCCCTCCACCAAGGCTCAGCTATTCAGATTGAGTCCCTTCAAGGAGAGAA 360
 QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
 DB 361 ACTATGCCCCCATTCAGAGCTGCTCTTCCAAATGTTAACTGACACAGATGGTGGG 420
 QY 141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
 DB 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGCAAGGCTTGGAC 480
 QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
 DB 481 ATCGATTTCAGAGAGAGTGGGCTTACATGGAATTCATGGTTCCACTCCCGGAAATC 540
 QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluAsnArgAlaAsnAsnValIleLysGly 200
 DB 541 TACGACCCGGAATTCATCTGCTGCAACAGAGAGCGTGCACAAACGATTAAGGCGC 600
 QY 201 ThrLysGlnGluGlnValGlnIleIleIleLysAspIleLysAlaPheLysGluAlaThr 220
 DB 601 ACAAGCAGAGAGCAATTCAGCAAAATCATCAAGACATCAAGCCGTTTAAAGAACCCACC 660
 QY 221 LysValAspLysValValLeuTyrThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
 DB 661 AAAGTGCAGAGGTGGTGTCTCTGCTGACTGCCCAACAGAGAGGTATAGCAATTTGGTT 720
 QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu 260
 DB 721 GTAGGCTTAAAGACCACTGAGAAATCTCTGGCTGCTGAGCAAGAAATGAGGCTGAG 780
 QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
 DB 781 ATTTCTCTTCCACCTTGTATGCAATTCGCTGTGTGATGAAATATTTCTTTCAATTAAT 840
 QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
 DB 841 GGAAGCCCTCAGAAACCTTTTGTACAGAGGCTGATGATCTTGCATCCCGAGAACTACT 900
 QY 301 LeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
 DB 901 TTGATTGGTGGAGATGACTTCAAGAGTGTCAAGCAAAATGAATATCTGTGTGGTGAAT 960
 QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
 DB 961 TTTCTTGTGGGGCTGGTATCAAGCCAACTATATAGTTAGTTACAAACATCTGGGAAAC 1020

QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
 DB 1021 AATGATGGTATGAAATCTTCGGCTCCCAAAACCTTCGCTCCAAAGGAAATCTTCCAAAGAC 1080
 QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
 DB 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTCTTATGAGCTGTGTGAACATCCC 1140
 QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
 DB 1141 GACCATGTGTTGTTATTAATGATGCTTACGTAGGAGATAGCAAGAGGCAATGAT 1200
 QY 401 GlnTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
 DB 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAAGAACACCATGTTTTCACAAACATAT 1260
 QY 421 GluAspSerLeuLeuAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
 DB 1261 GAGGATTCCTTTTACCTCTCTTATTAATCTTGAGCTTGTCTTCTGAGCTGAGC 1320
 QY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
 DB 1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGAAATTCACATTCACACCACTTGGCT 1380
 QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
 DB 1381 ACCATTCTCAGTATTCAGCAAGGCTCCTCTGTTTCCACCGGTACACACAGTGGTGAAT 1440
 QY 481 AlaLeuSerLysGlyAlaArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
 DB 1441 GCATTGTCAAAACAGCGTCAATGCTGGAAGAAACATATAGAGGCTTGTGTGATTGGCC 1500
 QY 501 ProGluAsnAspMetIleLeuGluTyrLys 510
 DB 1501 CCAGAGATATACATGATTCCTCAGATACAG 1530

RESULT 12
 ADS82003
 ID ADS82003 standard; cDNA; 1533 BP.
 XX
 AC ADS82003;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 4.
 XX
 KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
 KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
 XX
 OS Glycine max; line 29018JP03.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..1533
 FT /tag= a
 FT /product= "myo-inositol 1-phosphate synthase"
 XX
 PN US2003074685-A1.
 PD 17-APR-2003.
 PD
 XX 11-MAR-2002; 2002US-00025003.
 PF
 XX 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PI Hitz WD, Sebastian SA;
 XX
 DR WPI; 2004-639957/62.
 DR P-PsDB; ADS82004.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

XX Example 8, SEQ ID NO 13, 34pp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.

XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

XX Alignment Scores:
Pred. No.: 3,17e-250 Length: 1533
Score: 2600.00 Matches: 503
Percent Similarity: 99.41% Conservative: 4
Best Local Similarity: 98.63% Mismatches: 3
Query Match: 98.78% Indels: 0
DB: 13 Gaps: 0

US-10-718-952-2 (1-510) x ADS82003 (1-1533)

QY 1 MetPheIleGluAsnPhelYsValGluCyProAsnValIyetrThrgluThrgluile 20
DB 1 ATGTTATTCAGAAATTTTAAAGTAGAGAGTCCTAATGTGAAGTACCGAGACTGGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrgluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGCTGACAACTACGAAACCAACGAACTGTTCACAGAAACAGGATGGCACTAT 120
QY 41 GlnTPIleValIySerProIySerValIyTyrGluPhelYsThrAsnIleHisValPro 60
DB 121 CAGTGAATGTCTCAACCAATCCGTCACTCAATTTAAACCAACCCCATGTTTCCA 180
QY 61 LysLeuGluValMetLeuValGluTyrGluYAsnAsnGlySerThrLeuThrgluY 80
DB 181 AAATTGGGGGTGAGCTTGTGGGTGGGTGAAACAAGCTCTTACCTCCACCGGTGGT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTyrPalaThrLysAspIyAsileGlnIleAlaAsn 100
DB 241 GTTATTGCTTAAAGAGAGGCAATTCATGGGCTTAAAGAGCAAGATTCAACAGCAAT 300
QY 101 TyrPheGlySerLeuThrgluAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTTGGCTCCCTCCACCAAGCTCAGCTATTGAGTTGATCTTCCAGGAGAGGAA 360

QY 121 IleTyrAlaProPheIySerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCATTCAGAGTCTGCTTCCAATGTTAATCTCGACGACATTTGTGGG 420
QY 141 GlyTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGGATATCAGAAACATGAACTGGCTGATGCAATGGCCAAAGGTGTTGAGC 480
QY 161 IleAspLeuGlnYsGluLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTTCAGAACAGATTGAGGCTTACATGAAATCATGATCTCCACTCCCGGAAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleYsGly 200
DB 541 TAGACCCGGATTTATTGCTCTCCAAACAGAGGAGCGCCAAACAGTGAATTAAAGGCG 600
QY 201 ThrLysGlnGluGlnValGlnIleIleIleYsAspIleYsValPheYsGluAlaThr 220
DB 601 ACAAAGCAAGAGCAAGTTCCAGCAATCATCAAAAGCATCAAGCGTTTAAAGAAAGCCACC 660
QY 221 LysValAspLysValValIleuThrPThrAlaAsnThrgluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGCACAAAGTGTGTTCTGTGGACTGCCAACACAGAGGTTATGCAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuValAlaValAspArgAsnGluAlaGlu 260
DB 721 GTAGGCTTTAATGACACCATGAGAAATCTCTTGCTGTGACAGAAATAGGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCySerValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCTCTTCCACCTTGATGACATTCCTGTGTGAAGAAAGTTCCTTTCATTAAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAGAAACATTTTGTACAGGGCTGATTTGATCTTGCCATCCGAGGAACCT 900
QY 301 LeuIleGlyGlyAspAspPheIySerGlyGlnThrLysMetLysSerValIleuValAsp 320
DB 901 TTGATTTGGTGGAGATGACTTCAAGAGTGGTCAGCAAAATGAATCTGTGTGGTTGAT 960
QY 321 PheLeuValGlyValGlyIleYsProThrSerIleValSerTyrAsnHisIleuGlyAsn 340
DB 961 TTTCTTGTGGGGCTGTGATACAGGCACATCTAATAGTTAGTTAACCACTATGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIyGlnIleSerIySer 360
DB 1021 AATGATGATATCAATCTTCGGCTCCACAACTTCGGCTCCAAAGAAATCTCCAAAGAC 1080
QY 361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1081 AAGTTGTGTTACGATATGCTCAACAGCAATGCAATCTCTATGAGCTGTGTAACATCCC 1140
QY 381 AspHisValValIleIleYsTyrValProTyrValGlyAspSerIyAspArgAlaMetAsp 400
DB 1141 GACCAATGTTGTATTAAAGATATGCTTACGTAAGGGAATGAGAAAGACCAATGGAT 1200
QY 401 GlnTyrThrSerGlnIlePheMetGlyGlyLysSerThrIleValIleuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAAACACCATTTGTTTGCACAAACATGT 1260
QY 421 GlnAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleuLeuAlaGluLeuSer 440
DB 1261 GAGGAATCCCTTTTATGAGTCTCTATTAATCTTGAGACTTGCTCTTGCTGAGCTGAGC 1320
QY 441 ThrArgIleGluPhelYsAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGAAAAATTCATCTCATTCACCAAGTGGCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyTyrProValValAsn 480
DB 1381 ACCATTCTCAGCTATCTGACCAAGGCTCTCTGTGTTCCACCGGGTACACCAAGTGGTGAAT 1440

Qy	481		laleuserrvsglnarxalmetleuclutunentlemtaraglaaayeva glyleu aa	500
Db	1441	GCATTGCTAAAGCAGCGTGCATATGCTGGAAAAACATTAATGAGGCGTTGGTTGGATTGGCC	1500	
Qy	501		ProGlu u en a n e t le u g u t y l y s	510
Db	1501	CCAGAGATTAACATGATTCTCGAGTCAAG	1530	
RESULT 13				
ADQ14500				
ID	ADQ14500	standard; cDNA; 1533 BP.		
AC	ADQ14500;			
DT	23-SEP-2004	(first entry)		
DE	Mutant soybean myo-inositol 1-phosphate synthase cDNA #3.			
XX				
KW	Soybean; myo-inositol 1-phosphate synthase; gene; ss;			
KW	myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;			
KW	inorganic phosphate; mutant.			
XX				
OS	Glycine max.			
OS	Synthetic.			
XX				
FT	Key	Location/Qualifiers		
FT	CDS	1..1533		
FT		/*tag= a		
FT		/product= "Mutant soybean myo-inositol 1-phosphate		
FT		synthase #3"		
XX				
PN	US2004128713-A1.			
PD				
PD	01-JUL-2004.			
XX				
PF	21-NOV-2003; 2003US-00718952.			
XX				
PR	08-APR-1997; 97US-00835751.			
PR	07-APR-1998; 98WO-US006822.			
PR	26-APR-1999; 99US-00299315.			
PR	11-MAR-2002; 2002US-00025003.			
XX				
PA	(HITZ/) HITZ W D.			
PA	(SEBA/) SEBASTIAN S A.			
PA	(GRAC/) GRACE D J.			
PA	(SPRE/) STREIT L G.			
XX				
PI	Hitz WD, Sebastian SA, Grace DJ, Streit LG,			
XX				
DR	WPI; 2004-533135/51.			
DR	P-PSDB; ADQ14501.			
XX				
PT	New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,			
PT	useful for producing plants with decreased raffinose, stachyose, and			
PT	phytic acid and increased sucrose, leading to valuable and useful soybean			
PT	products.			
XX				
PS	Claim 10; SEQ ID NO 11; 48bp; English.			
CC				
CC	The invention relates to a nucleic acid fragment encoding a soybean myo-			
CC	inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate			
CC	synthase having a decreased capacity for the synthesis of myo-inositol 1-			
CC	phosphate. The invention also relates to a chimeric gene operably linked			
CC	to suitable regulatory sequences, where expression of the chimeric gene			
CC	results in a decrease in expression of an endogenous or native gene			
CC	encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant			
CC	comprising the chimeric gene, a method of making the soybean plant, a			
CC	seed of the soybean plant, a soy protein product derived from the			
CC	processing of soybean seeds, a method of making or producing a soy			
CC	protein product and a method of using a soybean plant homozygous for at			
CC	least one gene encoding a mutant myo-inositol 1-phosphate synthase having			
CC	decreased capacity for the synthesis of myo-inositol 1-phosphate. The			
CC	nucleic acid fragment and methods are useful for producing plants with			

[illegible]

```
OY 281 GlysSerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAGAAACCTTTTGTACAGGGCTGATGATCTTGCCATCGCGAAGAACT 900
OY 301 LeuIleGlyValAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTGGGAGAGTACTTCAAGAGTGGTCAGAACCAAAATGAAATGTTGTGGTGGAT 960
OY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 961 TTTCTTGGGGGGGGCTGGTATCAAGCCCAACATCTATAGTTACACACATCTGGGAAC 1020
OY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1021 AATGATGGTATGATATCTCGGCTCCACAACCTTCGCTCCAGAAGAAATCTCCAAGACC 1080
OY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlnHisPro 380
DB 1081 AACGTTGTTGACGATATGCTCAACAGCAATGCCATCTCTATGAGCCTGGTAAACATGCC 1140
OY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCATGTTGTTGTTATTAAGTATGTCCTTACGTAGGGGATACAGAGAGCCATGAT 1200
OY 401 GlnTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
DB 1201 GAGTACACTTCAAGATATTTCAATGGGTGGAAGAAGAACCATTTGTTGCACAACATGT 1260
OY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTTTAGCTCTCTATTAATCTTGGAATGGTGTCTTCTTCTCAAGCTGAGC 1320
OY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAGATATCAAGTTTAAAGCTAAATGAGGAAATTCACATCTTCCACCCAGTGGCT 1380
OY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB 1381 ACCATCTCAGCTATCTGACCAAGGCTCCTCTGCTTCCACCGGTTACACAGTGGGAT 1440
OY 481 AlaLeuSerLysGluArgAlaMetLeuGluAsnIleMetArgAlaGlyValGlyLeuAla 500
DB 1441 GCAATGTCAAAGCAGGTCGAAAGCTGGAACATATATAGAGGCTTGTGTGATTGGCC 1500
OY 501 ProGluAsnAsnMetIleLeuGlnTyrLys 510
DB 1501 CCAGAGATTAACATGATTTCTCCAGTACAG 1530

RESULT 14
ADS82001
ID ADS82001 standard; cDNA; 1533 BP.
XX
XX ADS82001;
XX
XX 18-NOV-2004 (first entry)
XX
XX Soybean myo-inositol 1-phosphate synthase cDNA mutant #2.
XX
XX DE Soybean, plant; myo-inositol 1-phosphate synthase; ss; gene; phylic acid;
XX KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
XX KW
XX OS Glycine max; line 29010CP01.
XX
XX
XX
XX FH
XX Key Location/Qualifiers
XX CDS 1..1533
XX FT /product= "myo-inositol 1-phosphate synthase"
XX FT replace(260,G)
XX FT /tag= b
XX
XX US2003074685-A1.
XX
```

```
PD 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
PF
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1996; 96WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
DR P-PSDB; ADS82002.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phylic acid and inorganic phosphate content of soybean seeds.
PS Claim 8; SEQ ID NO 11; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phylic
CC acid content of less than 14.5 micromol/g, a seed sucrose content of
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phylic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a mutant myo-
CC inositol 1-phosphate synthase.
XX
XX SO Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,59e-249 Length: 1533
Score: 2593.00 Matches: 502
Percent Similarity: 99.22% Conservative: 4
Best Local Similarity: 98.43% Mismatches: 4
Query Match: 98.52% Indels: 0
DB: 13 Gaps: 0

US-10-718-952-2 (1-510) x ADS82001 (1-1533)
OY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 1 ATGTTCAATCAGAAATTTTAAAGGTAGAGTCTTAATGTGAATACACCAAGCTAGATT 60
OY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGCTGATACACTACGAAACCAACCAACTTGTTCACGAGAACAGAAATGCACTAT 120
OY 41 GlnTrpIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
```

```

Db 121 CAGTGATGTCACCAACCAATCCGTCACTACCAATTTAAACCAACCACTGTTCCA 180
Qy 61 LysLeuGlyValMetLeuValGlyTPGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db 181 AAATTGGGGGTGATGCTTGGTGGGTGGGAAACACCGCTCTACCCCTCACCGGTGGT 240
Qy 81 ValIleAlaAsnArgGlyIleSerTPAlaThrLysAspLysIleGlnGlnAlaAsn 100
Db 241 GTTATTGCTAACAGAGAGACATTCATGGGCTACAAAGACAAAGATTCAACAAAGCAAT 300
Qy 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGln 120
Db 301 TACTTTGGCTCCCTCAACCAAGCTCAGCATTCGAGTTGATCTTCCACGGAGAGAA 360
Qy 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 361 ATCTATGCCCATTCACAGAGTCTGCTCCAAATGTTATCTCGACGACATGTGTTGGG 420
Qy 141 GYTTPAspIleSerMetMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 421 GGATGGGATATCGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTTTGAC 480
Qy 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyIle 180
Db 481 ATCGATTTGCAGAGACAGATTGAGGCTTACATGGAATCCATGGTTCCACTCCCGAATC 540
Qy 181 TyrAspProAspPheIleAlaAlaAsnGlnGlyLysArgAlaAsnValIleLysGly 200
Db 541 TACGACCCGATTTCAATGCTGCTCCAAACCAAGAGAGGTGCCAAACAGGATTAAGGGC 600
Qy 201 ThrLysGlnGlnGlnAlaGlnGlnIleIleLysAspIleLysValPheLysGlnAlaThr 220
Db 601 ACAAAACAGAGAGCAAGTCAAGCAATCATCAAAAGACATCAAGGCTTTAAGAGAGCACCC 660
Qy 221 LysValAspLysValValIleLeuTPThrAlaAsnThrGlyLysSerAsnLeuVal 240
Db 661 AAAGTGCACAGAGTGTGTTGCTCTGTGACCTGCCAACAGAGAGGATGATGCAATTTGGTT 720
Qy 241 ValGlyLeuAsnAspThrMetGlyAsnLeuLeuAlaValAspArgAsnGlnAlaGln 260
Db 721 GTAGGCTTAATGACACCATGAGAAATCTCTGGCTGTGGACAGAAATGAGCTGAG 780
Qy 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 781 ATTTCTCTCCACCTTGATGTCATGCTGTGTATGGAATAATGCTTTCATTAAAT 840
Qy 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 841 GGAAGCCCTCAGAACACTTTTGTACAGGGGCTGATGATCTTCCATCCGAGAGAACACT 900
Qy 301 LeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 901 TTGATTTGGTGAAGATGACTTCAAGAGTGTGTCAAGCAAAATGAATCTGTGTTGAT 960
Qy 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnIleLeuGlyAsn 340
Db 961 TTTCTTGGGGGCTGTGTATCAAGCCAAATCATATGATTAGTTAACACCACTCGGGAAC 1020
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyIleSerLysAsp 360
Db 1021 AATGATGATGATGATCTTCGGCTCCAAACCTTCCTCCAAAGAAATCTCCAAAGAG 1080
Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyGlnIleAsp 380
Db 1081 AACGTTGTTGACATATGTCACAGCAATGCAATGCAATCTCTATAGGCTGTGTAACATCCC 1140
Qy 381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db 1141 GACCAATGTTGTTATTAAGATATGCTTAAGTGGGATGCAAGAGAGCCATGAGAT 1200
Qy 401 GlnTyrThrSerGlnIlePheMetGlyLysSerThrIleValLeuHisAsnThrCys 420

```

```

Db 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAAGAACACCACTTTGTCACACATGT 1260
Qy 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGlnLeuSer 440
Db 1261 GAGGATTCCTCTTACTCTCTCTATATCTTGAGACTGTGCTTGTGCTGAGCTGAGC 1320
Qy 441 ThrArgIleGluPheLysAlaGlnLeuGlnGlyLysPheHisSerPheHisProValAla 460
Db 1321 ACTGATATCAGTTTAAAGCTGAAATGAGGAAATTCACATTCATTCACCACTTGTCT 1380
Qy 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db 1381 ACCATTCAGCATATGACCAAGGCTCTCTGTTCCACCGGCTACACCAAGTGTGAAT 1440
Qy 481 AlaLeuSerLysGlnArgAlaMetLeuGlnAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1441 GCATTCCTCAAGCAGCGTCAATGCTGGAACCATATATGAGGCTTGTGATTTGGCC 1500
Qy 501 ProGluAsnAsnMetIleLeuGlnTyrLys 510
Db 1501 CCAGAGATTAACATGATCTCGAGTACAAAG 1530

RESULT 15
ID AAX90402 standard; cDNA to mRNA; 1950 BP.
XX AAX90402;
AC AAX90402;
DT 24-SEP-1999 (first entry)
XX
DE Nicotiana paniculata INPS encoding cDNA.
XX Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;
XX water stress; resistance; de.
XX Nicotiana paniculata.
OS
XX JP11187879-A.
XX
XX 13-JUL-1999.
XX
XX 26-DEC-1997; 97JP-00359773.
XX
XX 26-DEC-1997; 97JP-00359773.
XX
PA (NINS) JAPAN TOBACCO INC.
XX
DR WPI: 1999-451546/38.
XX P-PSDB; AAY24477.
XX
PT New INPS gene derived from Nicotiana genus plant - useful for conferring
XX resistance to water stress to plants.
XX
PS Claim 2; Page 6-8; 8pp; Japanese.
XX
XX The present sequence encodes Nicotiana paniculata inositol monophosphate
CC synthase (INPS), designated NpINPS1. INPS can be used to confer water
CC stress resistance to a plant
XX
SQ Sequence 1950 BP; 570 A; 377 C; 425 G; 578 T; 0 U; 0 Other:

Alignment Scores:
Pred. No.: 2,93e-237 Length: 1950
Score: 2472.00 Matches: 470
Percent Similarity: 96.86 Conservative: 24
Best Local Similarity: 92.16 Mismatches: 16
Query Match: 93.928 Indels: 0
DB: Gaps: 0

US-10-718-952-2 (1-510) x AAX90402 (1-1950)
Qy 1 MetPheIleGluAsnPheLysValGlyCysProAsnValLysTyrThrGlnThrGlnIle 20

```

DB 92 ATGTTATTGAGAAATTTTAAGGTGAGACCCCAAGCTTAAGTACACCGAAAGTGAAT 151
QY 21 GlnSerValTyrAsnTyrGlnThrPhrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 152 CACTGCTGTATGATTAATCAAAACCACTGATGTTAGTTCAATGATGAGAAAATGGACATAT 211
QY 41 GlnTrpIleValIleAspProLysSerValIleTyrGluPheLysThrAsnIleHisValPro 60
DB 212 CAATGGACCGTCAGGCTAAGACTGTCAATATAGATTCAAGACTGATCTCATGTTCC 271
QY 61 LysLeuGluValMetLeuValGlyTyrGluYasnAsnGlySerThrLeuThrGlyGly 80
DB 212 AATATAGGGGTTATGCTTGTGATGGGGTGAACAAATGTTCAACCTTGCACCGGTGT 331
QY 81 ValIleIleAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnGlnIleAsn 100
DB 332 GTTATTGCTTACACGAGAAAGAAATTCATGGGCCCAAAAGATTAAGGTGCAACAGCAAT 391
QY 101 TyrPheGlySerLeuThrGlnIleAsnAlaIleArgValGlySerPheGlnGlyGlu 120
DB 392 TACTTGGCTCTCTTACTCAGGCTTCTACTATTCGAGTTGGGCTTTCAATGGAGAGAG 451
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 452 ACTTATGCTCCATTTAAAGCCCTCTTCCAAATGGTCAATCCAGATGACGTAGTGTGGA 511
QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 512 GGATGGGACATCAGCAGCAATGAATTTAGCAGATGCAATGGCCAGGCTTAAGGATTTGAT 571
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 572 ATTGATCTACAAAGAGAGTTGAGGCCCTTACATGGAATCTATGGTCCACTCCCTGGTATC 631
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 632 TATGACCTGATTTCTTGTCTGTACCAAGGGTCAAGTGCACGACCAACAGTGATCAAGGA 691
QY 201 ThrLysGlnGlnValGlnIleIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
DB 692 ACCAAGAAAGAAACAAATTCATCAATCATTAAGATTTAGGAGTTTAAGAAAGAAAC 751
QY 221 LysValAspLysValValIleuTyrPheAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 752 AAGTGACAGAGTGTAGTATGTTGACTGCTAACACTGAAAGATACAGTAATGTGTT 811
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluIleGlu 260
DB 812 GTTGAATTATGACCTATGAGAAACCTTTTGTCTGTGTGACAGAAATGAACTGAA 871
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 872 ATATCTCCTTCCACTTGTATGTATGTGCTGCTGCAATCTTGAAAATGTGCTTTTATTAAT 931
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 932 GGAAGCCCCAGAAACACCTTTGTCAGGACCTCATTTGATTTGGCCATCAAGAAACACA 991
QY 301 LeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB 992 TTGATTGGTGTGATACCTTTAAGAGTGTCAAAACCAAAATGAAATGACGTGTGTGAT 1051
QY 321 PheLeuValGlyValGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 1052 TTCCTTGTGGAGCTGTATTAAGCCAACTCAATGTGAGCTACAAACATTTGGGTAAC 1111
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1112 AATGATGAGATGAATCTGTCTGCCCTCAAACTTTCGGGTCAAAAGAGATCTGAAAAGT 1171
QY 361 AsnValValAspAspMetValAsnSerAlaAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1172 AATGTTGTGATGACATGTTTCAMAGCAATGCCATCTTATGAGCTGGAGACACCT 1231

QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1232 GACCATGTTGTGTGATTAAGATATGTGCATATGTGGAGACAGCAAGGGCAATGAT 1291
QY 401 GlnTyrThrSerGlnIlePheMetGlyLysSerThrIleValLeuHisAsnThrCys 420
DB 1292 GAGTACACATCTGAGATTTTCAATGGGGGAAAGAACACCATTTGTTTCCAAATACTGT 1351
QY 421 GluAspSerLeuLeuAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1352 GAGGATTCATTTAGCTGTCTCCATTAATTTGATTTGGTCTTGTGTAATCTCAGT 1411
QY 441 ThrArgIleGluPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
DB 1412 ACCGCATTCAGCTCAAGCTGAAGAGAGGGTAAATTCCACTCCCTCCACCCCGTGGCT 1471
QY 461 ThrIleLeuSerTyrLeuThrIleAlaProLeuValProProGlyThrProValValAsn 480
DB 1472 ACTATCTCAGCTACCTTACCAAGGCTCTCTGTGTACACACAGTACACAGGTGTGAT 1531
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1532 GCACTCTCAAGCAGAGGGCAATGCTTGAGAACTATTTGAGGGCTTGTGTTGACTTGA 1591
QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
DB 1592 CCAGAGAACACATGATTTCTGGAATACAAA 1621

Search completed: June 7, 2005, 17:47:34
Job time : 677.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 7, 2005, 17:22:56 ; Search time 219 Seconds
(without alignments)
3810.507 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632
Sequence: 1 MLENNKVCSPNKTETETI.....NIMRACVGLAPNNMILEYK 510

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 segs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=xlp
-Q=/cg2_1/USPTO.spool_P/US10718952/runat_06062005_173401_12868/app_query.fasta_1.1294
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=BITS -START=1 -END=1 -MATRIX=blomsu62 -TRANS=humand0.cdt
-LIST=45 -DOCALLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORMex -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10718952.0@CN_1.141.0@runat_06062005_173401_12868 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGIOLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cg2_6/prodata/1/ina/5A_COMB.seq:*
2: /cg2_6/prodata/1/ina/5B_COMB.seq:*
3: /cg2_6/prodata/1/ina/5A_COMB.seq:*
4: /cg2_6/prodata/1/ina/5B_COMB.seq:*
5: /cg2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cg2_6/prodata/1/ina/backfillseq.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2336	88.8	1931	3	US-09-118-442-10 Sequence 10, Appl
2	2336	88.8	1931	3	US-09-677-064-10 Sequence 10, Appl
3	2335	88.7	1939	4	US-09-727-628-1 Sequence 1, Appl
4	1813	68.9	3546	3	US-09-118-442-15 Sequence 15, Appl
5	1813	68.9	3546	3	US-09-677-064-15 Sequence 15, Appl
6	1812	68.8	3546	3	US-09-118-442-14 Sequence 14, Appl
7	1812	68.8	3546	3	US-09-677-064-14 Sequence 14, Appl
8	1382.5	52.5	1602	4	US-09-734-237B-72 Sequence 72, Appl
9	1382.5	52.5	1602	4	US-09-734-237B-74 Sequence 74, Appl
10	1376.5	52.3	1578	4	US-09-248-796A-3131 Sequence 3131, Ap
11	1220.5	46.4	77626	4	US-09-949-016-12608 Sequence 12608, A
12	1098.5	41.7	1231	3	US-09-397-787-34 Sequence 34, Appl

c	13	545	20.7	34316	4	US-09-902-540-1257	Sequence 1257, Ap
	14	541	20.6	1335	4	US-09-902-540-5285	Sequence 5285, Ap
	15	155	5.9	294	4	US-09-313-294A-4684	Sequence 4684, Ap
	16	142	5.4	42325	4	US-08-311-731A-131	Sequence 131, Ap
c	17	129.5	4.9	4403765	3	US-09-103-840A-2	Sequence 2, Appl
	18	128.5	4.9	4411529	3	US-09-103-840A-1	Sequence 1, Appl
c	19	115.5	4.4	7571	3	US-08-961-527-2	Sequence 2, Appl
	20	113.5	4.3	274	3	US-09-118-442-21	Sequence 21, Appl
	21	113.5	4.3	274	3	US-09-677-064-21	Sequence 21, Appl
	22	113.5	4.3	2190	4	US-09-107-433-2104	Sequence 2104, Ap
	23	113.5	4.3	2196	4	US-09-583-110-2059	Sequence 2059, Ap
	24	112.5	4.3	2529	4	US-09-489-039A-4399	Sequence 4399, Ap
	25	111.5	4.2	1359	4	US-09-248-796A-6542	Sequence 6542, Ap
	26	111.5	4.2	2193	1	US-09-014-897-1	Sequence 1, Appl
	27	111.5	4.2	2193	1	US-08-731-716-3	Sequence 3, Appl
	28	111.5	4.2	2193	1	US-08-731-716-1	Sequence 1, Appl
	29	111.5	4.2	2193	1	US-08-731-716-3	Sequence 3, Appl
	30	111	4.2	8287	4	US-09-023-655-1441	Sequence 1441, Ap
	31	109.5	4.2	2244	4	US-09-583-110-405	Sequence 405, Ap
	32	109.5	4.2	2256	4	US-09-107-433-772	Sequence 772, Ap
	33	109	4.1	3421	4	US-08-956-171B-337	Sequence 337, Ap
	34	109	4.1	3421	4	US-08-781-986A-337	Sequence 337, Ap
	35	109	4.1	8040	1	US-08-596-291-1	Sequence 1, Appl
	36	109	4.1	8040	1	US-09-100-804-1	Sequence 1, Appl
	37	109	4.1	8043	5	PCT-US94-09943-1	Sequence 1, Appl
	38	109	4.1	8119	3	US-09-290-640-45	Sequence 45, Appl
	39	109	4.1	8119	3	US-09-665-615B-45	Sequence 45, Appl
	40	108	4.1	1052	3	US-08-961-527-244	Sequence 244, Ap
	41	107.5	4.1	5892	4	US-09-583-110-2582	Sequence 2582, Ap
	42	107.5	4.1	5904	4	US-09-107-433-2280	Sequence 2280, Ap
	43	106.5	4.0	2337	4	US-09-489-039A-1284	Sequence 1284, Ap
	44	106	4.0	1520	4	US-09-543-681A-2705	Sequence 2705, Ap
	45	106	4.0	3729	4	US-09-107-532A-1587	Sequence 1587, Ap

ALIGNMENTS

RESULT 1
US-09-118-442-10
; Sequence 10, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Kang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-10
Alignment Scores:
Pred. No.: 1.05e-283
Score: 2336.00
Percent Similarity: 93.33%
Best Local Similarity: 87.25%
Query Match: 88.75%
DB: 3
Length: 1931
Matches: 445
Conservative: 31
Mismatch: 34
Indels: 0
Gaps: 0

US-10-718-952-2 (1-510) x US-09-118-442-10 (1-1931)

```

QY 1 MetPhe11e1e1uAsnPhelysVal1G1uCsProAsnVal1LysTyrThrGluThrGlu1le 20
Db 99 AAGTTCATCGAGAGCTTCGCGCTCGAGACCCCGCTGCGGTACGGCCGATGAGATTC 158
QY 21 GlnSerVal1TyrAsnTyrGluThrThrGluLeuVal1HiG1uAsnArgAsnG1YthrTyr 40
Db 159 GAGTCGAGATACCGGTACGACACGAGCTGCTGTCACGAGGACGAGGACGGCCCTCA 218
QY 41 GlnTP1e1e1Val1LysProLysSerVal1LysTyrGluPheLysThrAsn1Le1HiSVal1Pro 60
Db 219 CGCTGGGTGTCGCCGCCCAAGTCCGTCAGATCAACTTCGAGACCAAGCCCGCTCCC 278
QY 61 LysLeuG1uVal1MetLeuVal1G1YTPG1Yg1YAsnAsnG1YSerThrLeuThrG1Yg1Y 80
Db 279 AAGCTCGGGGTATGCTTGTTGGGTGGGAGGCAACACGGGTCCACGCTGACGGCTGGG 338
QY 81 Val11e1e1aAsnArgG1uG1Y1LysSerTPA1aThrLysAspLys1e1e1GlnG1n1aAsn 100
Db 339 GTTCATTCGCCAACAAGGAGGGAGTCTCATGCGGACCAAGGACAAAGTCCAGACCAAC 398
QY 101 TyrPheG1YSerLeuThrGln1aSer1a1e1ArgVal1G1YSerPheGlnG1YglnG1u 120
Db 399 TACTACGGCTCCCTCACCCAGCGCTCCACATCAGAGTCGCGCAGCTCAACAGGGAGAG 458
QY 121 11e1Yr1a1aProPheLysSerLeuLeuProMetVal1AsnProAspAsp11e1a1PheG1Y 140
Db 459 AACTATGCGCCCTTCACAGAGCTCTCTCCATAGTGAACCCAGACCAACATTTGTTCGGA 518
QY 141 G1YTPAsp11e1SerAsnMetAsnLeu1a1aAsp1aMet1a1Arg1a1aLysVal1PheAsp 160
Db 519 GGCTGGGACATTAAGCAACATGAACCTCGCGCATCATACCAAGGCCCAAGGTCTCGAT 578
QY 161 11eAspLeuGln1Yg1n1e1uArgProTyrMetG1uSerMetLeuProLeuProG1Y1le 180
Db 579 ATTGACCTGCAGAGAGAGCTCAGGCCCTCATGAGTCCATGATGCTCCCGTATC 638
QY 181 TyrAspProAspPhe11e1a1a1e1AsnGlnG1uArg1a1aAsnAsnVal11e1Ysg1Y 200
Db 639 TATGATCCGAGCTTCATCCGCGCTAACCGAGCTCTCGGCCAAGTGTCTCATCAAGGGC 698
QY 201 ThrLysGlnG1uGlnVal1GlnGln1e1e1LysAsp11e1LysAlaPheLysG1u1aThr 220
Db 699 ACCAAGAAGAACAAGGTGAGCAGATCATCAAGGATATCAGGAGCTTTAAGAGAAAGAAC 758
QY 221 LysVal1AspLysVal1Val1LeuTyrPThr1a1aAsnThrG1uArgTyrSerAsnLeuVal 240
Db 759 AAAGTCGACAAAGATAGTTGTGTGTGACTGCAAAACAAGTAAAGTATGCAATGTGTGC 818
QY 241 Val1G1YLeuAsnAspThrMetG1uAsnLeuLeu1a1a1Val1AspArgAsnG1u1aG1u 260
Db 819 GCTGGTCTAACGACACGATGAGAAATCTGAGCTGTGTGACAAAGAACAGAGGGAG 878
QY 261 11eSerProSerThrLeuTyr1a1e1a1e1aCyVal1MetG1uAsnVal1ProPhe11eAsn 280
Db 879 GTATCCCATCAACACTATATGCAATGCTGCTGTGATAGGAAGGGTGTGCCTTCATCAAT 938
QY 281 G1YSerProGlnAsnThrPheVal1ProG1YLeu11eAspLeu1a11e1a1ArgAsnThr 300
Db 939 GGGAGGCCCCAGAACACTTTGTGCTGGGCTGATGATCTGTCTATAAAAAACAATCTGC 998
QY 301 Leu11e1e1G1YAspAspPheLysSerG1YglnThrLysMetLysSerVal1LeuVal1aAsp 320
Db 999 TTGATTTGGTGTGACCACTTCAAGAGTGAACAACAACAAATGAATCTCTTGTGTCGAT 1058
QY 321 PheLeuVal1G1YAlaG1Y1Le1YsProThrSer11e1Val1SerTyrAsnHi1e1uG1YAsn 340
Db 1059 TTCCTTTGTTGTGCTGGAATTAAGCCCACTCATGTATAGCTACCAACACTTGGGAAAC 1118
QY 341 AsnAspG1YMetAsnLeuSer1a1aProGlnThrPheArgSer1Ysg1u1eSer1YsSer 360

```

```

Db 1119 AACGATGCGATGAACCTGCTGCCCCCTCMAACATTCAGGTCCAAAGAGATCTCCAAAGAC 1178
QY 361 AsnVal1Val1aAspAspMetVal1AsnSerAsn1a1e1e1uTyrGluProG1Yg1uHi1aPro 380
Db 1179 AACGTGTGATGACATGATGCTTCAGACAAATGCCATCTCTATGAGCCCGGCGAGATCCC 1238
QY 381 AspHiSVal1Val11e1YsTyrVal1ProTyrVal1G1YAspSerLysArg1a1aMetAsp 400
Db 1239 GATCATGTGTTGTATCATCAAGATATGTCGTAAGCGGAGACAGACAAAGAGGCTATGAGAC 1298
QY 401 G1uTyrThrSerG1u11ePheMetG1Yg1YLysSerThr11eVal1LeuHi1aAsnThrCys 420
Db 1299 GAGTACACCTCAGAAATTTATATGAGCGGCAAGAACACCATCTGTGCTCACAACACCTGT 1358
QY 421 G1uAspSerLeuLeu1a1a1aPro11e1e1e1uAspLeuVal1LeuLeu1a1e1u1e1uSer 440
Db 1359 GAGGACTGCTCTCGCCGCACTTATCTTATCTGTGCTGCTTGGCTGAGCTCAGC 1418
QY 441 ThrArg11e1e1uPheLysAlaG1uAsnG1uG1YLysPheHi1aSerPheHi1aProVal1a 460
Db 1419 ACCAGATCCAGCTGAAGCTGAGGAGAGGACAAATTCACACTCTTCACCCGGTGGCC 1478
QY 461 Thr11e1e1uSerTyrLeuThrLysAlaProLeuVal1ProProG1YThrProVal1a1aAsn 480
Db 1479 ACCATCTTGAAGTACTTCAACAAAGCACCCCTGTCTCCCTGCGACACGGGTGTGAAC 1538
QY 481 AlaLeuSerLysG1YArg1a1aMetLeuG1uAsn11eMetArg1a1aCyVal1G1YLeu1a 500
Db 1539 GCTTGGCCCAACAGAGGCGCATGCTGAGAAACATCATAGAGGCTCGCTGGCTGGCC 1598
QY 501 ProG1uAsnAsnMet11e1e1uG1uTyrLys 510
Db 1599 CCAGAGAACACATGATCTTGGAGTACAAAG 1628

```

RESULT 2

US-09-677-064-10
 / Sequence 10, Application US/09677064
 / Patent No. 6291224

GENERAL INFORMATION:

/ APPLICANT: Martino-Galt, Susan J.
 / APPLICANT: Beach, Larry R.
 / TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
 / FILE REFERENCE: 0706D
 / CURRENT APPLICATION NUMBER: US/09/677,064
 / CURRENT FILING DATE: 2000-09-29
 / PRIOR APPLICATION NUMBER: 60/055,446
 / PRIOR FILING DATE: 1997-08-11
 / PRIOR APPLICATION NUMBER: 60/055,526
 / PRIOR FILING DATE: 1997-08-08
 / PRIOR APPLICATION NUMBER: 60/053,944
 / PRIOR FILING DATE: 1997-07-28
 / PRIOR APPLICATION NUMBER: 09/118,442
 / NUMBER OF SEQ ID NOS: 31
 / SOFTWARE: FastSeq for Windows Version 3.0
 / SEQ ID NO 10
 / LENGTH: 1931
 / TYPE: DNA
 / ORGANISM: Zea mays
 / US-09-677-064-10

Alignment Scores:

Pred. No.:	Length:
1,05e-283	1931
Score: 2336.00	Matches: 445
Percent Similarity: 93.33%	Conservative: 31
Best Local Similarity: 87.25%	Mismatches: 34
Query Match: 88.75%	Indels: 0
DB: 3	Gaps: 0

US-10-718-952-2 (1-510) x US-09-677-064-10 (1-1931)

```

QY 1 MetPhe11eGluAanPhe1yValGluCySProAanVal1yFyThrGluThrGlu1e 20
Db 99 ATGTTTCATGAGAGAGCTTCGCGCTCGAGAGAGCCCACTGCGGTAACGGCCGATGAGATC 158
QY 21 GlnSerVal1TyraAnTyrgluThrThrgluLeuVal1H1eGluAanArgAnGlyThrTy 40
Db 159 GAGTCGAGATACCGGATGACACAGACGAGAGCTGTACACAGAGGCAAGAGCGCGCTCA 218
QY 41 GlnTP11eVal1ySerPro1ySerVal1yFyThrGluPhe1yThraAn1eH1eValPro 60
Db 219 CCTGGGTGCTCGCCCAAGTCCGTCAGTACATTCCTCGGACCAAGACCGCGCTCCC 278
QY 61 LysLeuGlyValMetLeuVal1GlyTPG1yGlyAanAnGlySerThreThrThrgly 80
Db 279 AAGCTCGGGGTGATGCTTGTGGGTGGGAGGAGCAACAGGCTCCAGCTGACGCTGG 338
QY 81 Val11eAlaAanArgGluGly11eSerTP1a1aThraLysAsp1yA11eGln1a1aAn 100
Db 339 GTCATTCGCAACAGGAGGGATCTCATGGGCGACCAAGACAGAGGTGACGAGCCAAC 398
QY 101 TyrPheGlySerLeuThrGlnAlaSerAla11eArgVal1GlySerPheGln1yGlu 120
Db 399 TACTACGGCTCCCTCCACCGCTCCACATCAGAGTGGCGCATCAACGCGGAGAG 458
QY 121 11eTyra1aProPhe1ySerLeuLeuProMetVal1AanProAsp11eVal1PheGly 140
Db 459 ATCTATGCGCCGTTCAAGAGCTCTCTCCATAGTAACCCACACACATGTGTTGGA 518
QY 141 GlyTPAsp11eSerAanMetAanLeu1aAspAlaMetAlaArgAlaLysVal1PheAsp 160
Db 519 GGCTGGGACATTAGCAACATGAACCTGGCGCATGACCAAGGCGCAAGGTGCTGGAT 578
QY 161 11eAspLeuGln1yGlnLeuArgProTyrmGluSerMetLeuProLeuProGly11e 180
Db 579 ATTTGACCTGCACAAAGAGCTCAGGCCCTCATGAGATCATGTGCTCCGCTATC 638
QY 181 TyraAspProAspPhe11eAla1aAnGln1yGluArgAlaAanAanVal11eLysGly 200
Db 639 TATGATCCGGAATTCCTCGCGCTAACCGAGGCTCTCGGCGCAACGTGTCATCAAGGGC 698
QY 201 ThrLysGlnGluGlnVal1Gln11e11eLysAsp11eLysAlaPhe1ySerGlu1a1aThr 220
Db 699 ACCAAGAAAGAACAGGTGAGGAGCATCATCAAGATATCAGAGGATTTAAGAGAGAAC 758
QY 221 LysValaAspLysValaVal1LeuTPThra1aAnThrGluArgTySerAanLeuVal 240
Db 759 AAAGTGACAAAGATAGTGTGTGTGCTGCTGCAACACATGAAGGTATGACATGTGCTC 818
QY 241 ValGlyLeuAanAspThrMetGluAanLeuVal1a1aValaAspArgAnGlu1aGlu 260
Db 819 GCTGGTCTCAACGACGATGAGAACTTCTGCTGCTGACCAAGAACGAGCGGAG 878
QY 261 11eSerProSerThreThrTyra1a11eAlaCyValMetGluAanVal1ProPhe11eAn 280
Db 879 GATATCACCATCAACATATATGCTGCTGTGCTATGGAAGGGGTGCGCTTCATCAAT 938
QY 281 GlySerProGlnAanThrPheVal1ProGlyLeu11eAspLeuAla11eAlaArgAnThr 300
Db 939 GGGAGGCCCCCAACACCTTTGTGCTGGGCTGATGATCTTGTATTAACCAACTGC 998
QY 301 Leu11eGly1yAspAspPhe1ySerGly1nThyMetLysSerVal1LeuValaAsp 320
Db 999 TTGATTTGGTGGTGCACATTCACAGGTGACACACAAATGAATCTGCTTGGTCGAT 1058
QY 321 PheLeuValaGlyAlaGly11eLysProThrSer11eValaSerTyraAn11eLysGlyAan 340
Db 1059 TTCCTTTGGTGGTGCATTAAGCCCACTCAATGTGATGCTACACCACTTGGGAAC 1118
QY 341 AaAnAspGlyMetAanLeuSerAlaProGlnThrPheArgSer1yGlu11eSer1ySer 360
Db 1119 AACGATGGGATGAACCTGTCTGCTCCCTCAAACTTCAGGTCCCAAGAGATCTTCAAGAGC 1178
QY 361 AaAnVala1aAspAspMetValaAnSerAanAla11eLeuTyrgluProGlyGluH1ePro 380

```

```

Db 1179 AAGTGTGATGACATGCTTCGAGCATGCAATGCATCTCTATAGAGCCGCGGAGCATGCC 1238
QY 381 AspHisValaVal11eLysTyra1aProTyraValaGlyAspSer1yAra1aMetAap 400
Db 1239 GATCATGTGTTGTCTCAAGATATGTGCGGTACGTGGAGACAGCAAGAGGCTATGAGAC 1298
QY 401 GlyTyrrThreSerGlu11ePheMetGly1yLysSerThra1aVal1LeuH1eAnThrCyS 420
Db 1299 GAGTACACTTCAGAGATCTTCAAGGCGGACCAAGAACACATGTGTGACACACCTGT 1358
QY 421 GluAspSerLeuLeuAla1aPro11e11eLeuAspLeuVal1LeuLeuSer 440
Db 1359 GAGGACTCCCTCTCGCCGACCTATCATCTTGATCTGTGCTCTTGGCTAGCTCAGC 1418
QY 441 ThrArg11eGluPhe1yAlaGluAanGly1yLysPheH1eSerPheH1eProVala1a 460
Db 1419 ACCAGGATTCACGTGAACCTGAGGAGAGACAAATTCACATCTTCCACCGCGTGGCC 1478
QY 461 Thr11eLeuSerTyrrThrLysAlaProLeuVala1ProProGly1yThrProVala1aAn 480
Db 1479 ACCATCTTGAATTAATTCACCAAGGCACTGTTCCCTGGCGACACCGGTGTGAAC 1538
QY 481 AlaLeuSer1yGlnArgAlaMetLeuGluAan11eMetArgAlaCyVala1GlyLeuAla 500
Db 1539 GCTTGGCCCAACAGGCGCATGCTGTGAGAACATCATAGAGGCTGCTGGGCTGGCC 1598
QY 501 ProGluAanMet11eLeng1yTyrr1yS 510
Db 1599 CCAGAGAACACATGATCTTGAATGACAG 1628

RESULT 3
US-09-727-628-1
; Sequence 1, Application US/09727628
; Patent No. 6791013
; GENERAL INFORMATION:
; APPLICANT: Armstrong, Katherine
; APPLICANT: Hey, Timothy D
; APPLICANT: Folkerts, Otto
; APPLICANT: Smith, Kelley A
; APPLICANT: Hopkins, Nicole L
; TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
; FILE REFERENCE: 50597
; CURRENT APPLICATION NUMBER: US/09/727,628
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/168,612
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1959
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (137)..(1699)
; US-09-727-628-1

Alignment Scores:
Pred. No.: 1,44e-263 Length: 1959
Score: 2335.00 Matches: 447
Percent Similarity: 93.33% Conservative: 29
Best Local Similarity: 87.65% Mismatches: 34
Query Match: 88.72% Indels: 0
Gaps: 0

US-10-718-952-2 (1-510) x US-09-727-628-1 (1-1959)
QY 1 MetPhe11eGluAanPhe1yValGluCySProAanVal1yFyThrGluThrGlu1e 20
Db 137 ATGTTTCATGAGAGAGCTTCGCGCTCGAGAGAGCCCACTGCGGTAACGGCCGATGAGATC 196
QY 21 GlnSerVal1TyraAnTyrgluThrThrgluLeuVal1H1eGluAanArgAnGlyThrTy 40

```

```

Db      197 GAGTCGAGACTACCGGTACACAGACGAGAGCTGTGTCACGAGAGCCAAAGAGCGGCGCTCC 256
      41 GINTPILLeValLySerProLySerValLySerTyrGluPhelThraSnLLeHisValPro 60
      257 CGCTGGGTGCTCGCCCAAGTCCGTCAAGTCAACTTCGAGCAAGCAGCCGCGGCTCCC 316
Qy      61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrcGlyGly 80
Db      317 AAGCTCGGGGTGATGCTTGTGGGGTGGGAGGCAACAGGGTCCACGCTGACGGGTGGG 376
Qy      81 ValLLeaLAsnArgGlyGlyLysSerTyrPalArthrLysAspLysLLeGlnGlnLAsn 100
Db      377 GTCATATGCCAGCAGGAGGAGGATCTCATGGCGCACCAAGACAGAGGTGACGAGCAAGCAAC 436
Qy      101 TyrPhelGlySerLeuThrcLnaLAsnAlaLeaGlyValGlySerPhelGlnGlyGln 120
Db      437 TACTACGGGTCCCTCCAGGCTTCACCATCAGAGTACGAGCTCAACAGCGGAGGAG 496
Qy      121 LLeTyrAlaProPhelLySerLeuLeuProMetValAsnProAspAspLLeValPheGly 140
Db      497 ATATATGCCCGCTTCAAGAGCTCTCTACCAAGTGAACCAAGACGACCTTGTGTTGA 556
Qy      141 GlyTyrAspLLeSerAsnMetAsnLeuLAsnAlaMetAlaArgAlaLysValPheAsp 160
Db      557 GCGTGGACATCAGCAGCATGAACCTGGCAGATGCGATGACCAAGGCGCAAGGTGCTGAC 616
Qy      161 LLeAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyLLe 180
Db      617 ATTGACCTTCAGAGAGCTCAGGCTTCATGAGATGATGATGATGATGATGATGATGATG 676
Qy      181 TyrAspProAspPheLLeaLAsnGlnGlnLLeLysAspLLeLysValLLeLysGly 200
Db      677 TATGATCCCGGACTTCTCCCGCTTACCAAGGCTCTCGTCCCAACAATGTCATCAAGGCG 736
Qy      201 ThrLysGlnGlnLValGlnGlnLLeLysAspLLeLysValLLeLysGlnLLeLys 220
Db      737 ACCAAGAAAGAACAGGTGAGCATCATCAAAAGATATCAGGAGGATTAAAGGAGAAAC 796
Qy      221 LysValAspLysValValLeuTyrPThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db      797 AAGGTGAGAACAGGTATGTGTGTGCTGACTGCAACACATGAAAGTACAGCAATGTATGT 856
Qy      241 ValGlyLLeuAsnAspThrMetGluAsnLeuLLeaLValAspArgAsnGlnLLeLys 260
Db      857 GCTGGTCTCACGACACATGAGATCTGTGTGCTGTGTGCAAGAACAGAGGCGGAG 916
Qy      261 LLeSerProSerThrLeuTyrAlaLLeaLysValMetGluAsnValProPheLLea 280
Db      917 ATCTCCGCATCAACACTATATGCTATGCTGTGTGACGAGGCGGTGCTCATCAAT 976
Qy      281 GlySerProGluAsnThrPheValProGlyLLeuLLeaLysValLLeaLArgAsnThr 300
Db      977 GGGAGGCCCCAGAACCTTTGTGCTGGGCTGATGATCTTGTCAAGAACCAACTGCG 1036
Qy      301 LLeuLLeGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db      1037 CTGATCGGTGTGACACACTTCAAGAGTGGGACCAAGATGAATCGGTCTGTGAT 1096
Qy      321 PheLeuValGlyValGlyLLeuLysProThrSerLLeValSerTyrAsnLLeuLys 340
Db      1097 TTTCTTGTGTGCTGGAATTAAGCCACCTCGATGTGTGATCAACCACTGGGAAAC 1156
Qy      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysLLeuLysSer 360
Db      1157 AACGAGCGGATGAACCTGTCTGCTCCCTCAAAACATTCAGGCTCAAGAGATCTCAAGAGC 1216
Qy      361 AsnValValAspAspMetValAsnSerAsnAlaLLeuTyrGluProGlyGlnLLeuLys 380
Db      1217 AACGTGTGTGATACATGCTCTCAAGCAATGCAATCTCTATGTGGCCCGGCGAGCATCCC 1276
Qy      381 AspHisValValValLLeuLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400

```

```

Db      1277 GATCATGTTGTTCATCAAGTATGTGCCGTATGTGGAGACAGTAAAGGCTATGAC 1336
Qy      401 GluTyrThrSerGluLLePheMetGlyLysSerThrLLeValLeuHisAsnThrCys 420
Db      1337 GAGTACACATCAGAGATCTTCAATGGGCGGCAAGACACATGCTGTGCACAAACCTGCG 1396
Qy      421 GluAspSerLeuLeuAlaLLeaProLLeLLeuAspLeuValLeuLeuLLeuLLeuLLeu 440
Db      1397 GAGGACTGCTCTCGCGCGACCGATCATCTCGATCTGTGCTCTGCTGCTGAGCTCAGC 1456
Qy      441 ThrArgLLeGluPhelValAlaGluAsnGlnLysPheHisSerPheHisProValAla 460
Db      1457 ACCAGATATCCAGTTAAACCTGAGGGAACGACAAAGTTCACCTTCCACCGGAGGCGC 1516
Qy      461 ThrLLeuSerTyrLLeuThrLysAlaProLeuValProProGlyThrProValAla 480
Db      1517 ACCATCTTACCTTACCTTACCTTACCAAGGACACATCGTTCACCGGACACCGGTGTGAAC 1576
Qy      481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnLLeuMetArgAlaCysValGlyLeuAla 500
Db      1577 GCTCTTGCAAGCAGAGGCGATGCTGTGGAGAACATCATGAGGCTTGCGTGGCTGCGC 1636
Qy      501 ProGluAsnMetLLeuGlnLys 510
Db      1637 CCAGAGAACCAATCATATCTGTGATCAAG 1666

RESULT 4
US-09-118-442-15
/ Sequence 15, Application US/09118442B
/ Patent No. 6197361
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ APPLICANT: Wang, Xun
/ TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
/ FILE REFERENCE: 0706
/ CURRENT APPLICATION NUMBER: US/09/118,442B
/ EARLIER FILING DATE: 1998-07-17
/ EARLIER APPLICATION NUMBER: 60/055,446
/ EARLIER FILING DATE: 1997-08-11
/ EARLIER APPLICATION NUMBER: 60/055,526
/ EARLIER FILING DATE: 1997-08-08
/ EARLIER APPLICATION NUMBER: 60/053,944
/ EARLIER FILING DATE: 1997-07-28
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 15
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-118-442-15

Alignment Scores:
Pred. No.: 2,71e-217 Length: 3546
Score: 1813.00 Matches: 443
Percent Similarity: 40.86% Conservative: 31
Best Local Similarity: 38.19% Mismatches: 36
Query Match: 68.88% Indels: 652
DB: 3 Gaps: 9

US-10-718-952-2 (1-510) x US-09-118-442-15 (1-3546)
Qy      1 MetPheLLeGluAsnPhelLysValGluCysProAsnValLysTyrThrGluThrcGluLLe 20
Db      60 ATGTTTCATCGAAGAGCTTCCGCTCGAGAGCCCACTGCTGCTGAGCGCCCGAGCGAGATC 119
Qy      21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db      120 GAGTCGAGTACCGGTGACACAGACGAGGCTGTGTACAGAGGCGCAAGAGCGCGCTCA 179

```


QY 41 GlnTrpIleValIysPheProIysSerValIysIleThrGluPheLysPheThrAsnIleHisValPro 60
Db 180 CGCTGGGTGTCGCCCAAGTCCGTCAGAGTACAACTTCGACGACCAAGCCGCCGCC 239
QY 61 LysLeuGly----- 63
Db 240 AAGCTCGGGTATGATACGAGTACGAGCGGCCCTACCTCAGCTCTGTGAAACCTCTCTCTC 299
QY 63 ----- 63
Db 300 CCGTGTCTCAGTCAAAATCTCCGTCGAGATCACTGGTCGCGCTTCCCTCTTAATCTTAA 359
QY 63 ----- 63
Db 360 TGAATAATTCTACTGCTTGGCTTGAAGACGAAACCGTCGTAATTGTGTGACAGCTACGCAAC 419
QY 63 ----- 63
Db 420 ACTTGCCCATCCGGATGCGTCAAAATGAGCTCGATTGAAATTGATTCGATGGTGCCTT 479
QY 63 ----- 63
Db 480 TTCCATATTTCGATCATCCCTCGCTACTGTGCAATGATTACAGAAACGTCCTTCTCTC 539
QY 63 ----- 63
Db 540 TGAATTCTGTCTTAGCGTTTGTCTGTCTGTGACGTGACGTGATCAATTGTTCATGTA 599
QY 63 ----- 63
Db 600 AGATCAAAATTCACGACGAGGACGATGACGACGACAGACAGAACTCATACGCTAGCAATTGA 659
QY 63 ----- 63
Db 660 TACTAGATTACTGCAAGTGTGATACGCGCATCTGCATCTGACCCCTTTGTT 719
QY 63 ----- 63
Db 720 TAATTCTGTCTTCTATGATGATTGCTTACGTGACGCTGTTGTGTATGATGTCAGGC 779
QY 63 ----- 63
Db 780 TGTACGCGCTTGTCTCTGTGACGAGATGATGCCAACTTTCTGTCTGTGTGTGACAG 839
QY 64 ValMetLeuValGlyTrpGlyIleAsnAsnGlySerThrLeuThrGlyValIleAla 83
Db 840 GTGATGCTTGTGGGGTGGGAGGCAACAGGGTCCACGCTGACGCGTGGGTCAATTGCC 899
QY 84 AsnArgGlu----- 86
Db 900 ACGAGGAGTGAAGTACTTAATTGTCTTATGCTTTCGTTCCGTTGTTTCAGTTATTA 959
QY 87 ----- 92
Db 960 ATGGCTGACAGAACTGAATTTTGTGTGGCTGTTTCAGGGGATCTCATGGCCGAC 1019
QY 93 LysAspLysIleGlnGlnIleAsnIlePheGlySerLeuThrGlnIleAsnAlaIleArg 112
Db 1020 AAGGACCAAGGTGACGACCAACTACTACGGCTC-CTCACCCAGGCTTCACCATCAGA 1078
QY 113 ValGlySerPheGlnGlyGluGluIleIleIleIleIleIleIleIleIleIleIleIleIle 131
Db 1079 GTCGGAGGTACACAGGGAGAGATCTATGCGCTTCAGAGCCTCTCTCCCATGTGA 1138
QY 131 ----- 131
Db 1139 ATCTATTATGACTGATTAATCTCTCTTTTACTGAACCAACATATACAAAG 1198
QY 131 ----- 131
Db 1199 CATATTCCGTAAGGTCTAGTGTATTAAGTGAACCTGTCTTTCAGGCGAGTGTTC 1258
QY 131 ----- 131

Db 1259 TCAAGTAAACGGAATGTTAATCATTTGGGTGAAAAACAAGTTCATTTTGTGAAG 1318
QY 131 ----- 131
Db 1319 GAATGTTAACTTAGCATATGAAAAAGGAGACATTGTAAAGAAAGTGTGAACAATC 1378
QY 131 ----- 131
Db 1379 GACTGGTCTGCATGTTGTGATCTTACTTGAAGTCAAAAGTTCGTGTGTAAGTCAA 1438
QY 131 ----- 131
Db 1439 AGGTCCAGCATTTTGAATTAACCTGTGACGATATTGACATGTGTCTAAGTGTGCA 1498
QY 131 ----- 131
Db 1499 GATTGCAAGACTCGGTTGTTATCTTTCATGACCAAGTGTGAACGTGTTTCA 1558
QY 132 -ValAsnProAspAspIleValPheGlyGlyTrpAspIleSerAsnMetAsnLeuAlaAs 151
Db 1559 GGTGAACCCAGACGACATTGTGTTCGAGGCTGGAGCATTTAGCAACATGAACCTGCGCA 1618
QY 151 PalMetAlaArgAlaIysValPheAspIleAspLeuGlnLysGlnLeuArgProTyrMet 171
Db 1619 CTCATGACCAAGGCGCAAGTCTGTGATATTAACCTGCAAGAGCACTCAGGCCCTTACAT 1678
QY 171 GluSerMetLeuProLeu-ProGlyIleTyrAspProAspPheIleAlaAsnGlnG 191
Db 1679 GGAATCATGATGTCACCTTCCCGGATCTATGATCCGACCTTACCGGGCTAACAG 1738
QY 191 IugIuArgAlaAsnAsnValIleIleGlyThrLysGlnGlnValGlnGlnIleIleLeu 211
Db 1739 GCTCTCGCCCAACAGTGCATCAAGGCGACCAAGAAAGACAGGGAGAGCATCATCA 1798
QY 211 ValAspIle----- 213
Db 1799 AGGATATCAGGTATATGATATGATATGCTAACGTGCTGTAAAGTGCACCCAGTC 1858
QY 213 ----- 213
Db 1859 CAACCTAAACAAATTAATTAATGATTAATAATTGGTAATATATACATATCAGAACAT 1918
QY 213 ----- 213
Db 1919 ATGTGTTAACCGGTGACCTTAGAAGTCTGATGATGTTGGACAATTGACATTCGATA 1978
QY 213 ----- 213
Db 1979 TACAGTACCGCTCACTTGCATGAGAGACTCCACAAAGAACTAATAACTGAAAGCTTAA 2038
QY 213 ----- 213
Db 2039 GCAACTATTGTAAGTATGATGATTTGTGTGAGACATGTTGAAGATCTAGATTAACT 2098
QY 213 ----- 213
Db 2099 GGTGAAGAAATATGTTCACTAGCATAGTAATCATTAACAAGCTATGCTTATGTA 2158
QY 214 -----LysAlaPheLysGlnAlaThrLysValAspLysValValIleuTrpT 230
Db 2159 GCTTAATGAACAGGAGTTTAAAGAGAAAGAACAAAGTGAAGAATGTTGTGTGGA 2218
QY 2230 hTrAlAsnThrGluArgTyrSerAsnLeuValIleGlyLeuAsnAspThrMetGluAsnL 250
Db 2219 CTGCAAAACACTGAAGAGTATACATGTGTGCGTGTCTCAACACACGATGGAATC 2278
QY 250 euLeuAlaIleValAspArgAsnGluAlaGluIleSerProSerThrLeuTyrAlaIleAla 270
Db 2279 TACTGGCATCTGTGACAAAGACGAGCGAGGTATCAACATCAACACTATATGTCATTG 2338
QY 270 LaCysValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValPro- 289

```
Db 2339 CCTGTGTCATGAGAGGGGTGCCCTTCATCATGAGAGCCCGCAACACCTTTGTGCTG 2398
Qy 289 ----- 289
Db 2399 GTGGCGTGTGGTGTGTTGCAAAAGCTTCATGCTGTTCATTTCTGTCCAAAGTTTC 2458
Qy 290 -----GlyLeu1le 292
Db 2459 ATGCTGTGTATTTCCGTTCCAGGCTTATTTATACCTGTGTGCAATGTTCCGAGGGCTGAT 2518
Qy 293 AspLeuAla1leAlaArgAsnThrLeu1leGlyValAspAspPheLysSerGlyGlnThr 312
Db 2519 GATCTGTGTATATAAAACACCTGCTGATTTGGTGGAGAGACTTCAAGAGTGGACAGACC 2578
Qy 313 LysMetLysSerValLeuValAspPheLeuValGlyValGlyLeu1leLys----- 328
Db 2579 AAGATGCAATCTGTCTGTGTCATTTCTTTGTTGGTGTGGAAT-AAAAGTGGAAACCTA 2637
Qy 328 ----- 328
Db 2638 GATATCTCTCTCTATTAAGATGAGAGTGTGTTTGGCAATGACGTTATTGCAATTAATC 2697
Qy 329 -----ProThrSer1leValSer1leValSer1leValSer1leValSer1leValSer 340
Db 2698 TTCTATATTATTTTCATTTTCATGACAGCCCACTCAATGTAGCTACACCACTGGGAAAC 2757
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSer1leValSer1leValSer 360
Db 2758 AACGATGGCATGACCTGTCTGCTCCCTTCAAACTTCAGTCCAGAGAGATCTCCAGAGAC 2817
Qy 361 AsnValValAspAspMetValAsnSerAsnAla1leLeu1leGlyLeuProGlyLysIlePro 380
Db 2818 AACGTGTGATGACATGCTGTGACAGCAATGACCATCTCTATAGAGCCCGGACGATCCCC 2877
Qy 381 AspHisValValVal1leLys----- 387
Db 2878 GATCATGTGTGTGTTCATCAAGGTCTGTGTAGCTGATCTTCACTCGTTAAAGTTGACAT 2937
Qy 388 -----TyrValPro1leVal 392
Db 2938 ATGCAGAGCAGATTTCATTGAAACTGTGCATCTTTGTGACAGATATGTCCGATACG 2997
Qy 393 GlyAspSer1leValArg1leMetAspGlyLysThrSerGly1lePheMetGlyLysSer 412
Db 2998 GAGAGACAGCAAGGGGCTATGACGAGTACACCTCAGAGATCTTCATGCGGCAAGAAC 3057
Qy 413 Thr1leValLeuHisAsnThrCysGluAspSer1leValAlaPro1le1leLeuAsp 432
Db 3058 ACCATGTGCTGCACAACTGTGTGAGACTGTGCTCTGCGCGCACCTTATCATCTTGAT 3117
Qy 433 LeuValLeuValAlaGluLeuSerThrArg1leGlyPheLysAlaGluAsnGlu----- 450
Db 3118 CTGGTCTCTTGTGCTAGCTCAGACCAAGATCCACTGAAAGCTGAGGAGAGATGA 3177
Qy 450 ----- 450
Db 3178 GCCCCCAAGATTAACCTGAAGACGCTGACGCTAGTATATACACTTTTAATA 3237
Qy 451 -----GlyLysPheHisSerPheHisProValAlaThr1le 462
Db 3238 CCTTGTGTGTCTCTCTTATGACAGCAAAATTCATCTTCCACCCGGGTGGCAACATC 3297
Qy 463 LeuSer1leLeuThr1leLysAlaProLeu----- 471
Db 3298 CTGAGCTACTCTCAAGGCAACCCCTGTGTAAGCTTTTCTCTGCATCCCGGATCACTG 3357
Qy 472 -----Val 472
Db 3358 CACTGGCTTTTGGCTTCAATCCAGCACTGATGCTCTCTCGAAACCTGAACACAGATT 3417
Qy 473 ProProGly1leThrProValValAsnAlaLeuSer1leValArg1leMetLeuGluAsn1le 492
Db 3418 CCCCCTGGACACCCGGTGTGAACGCTTGGCCAAAGCAGACGGCGATGCTGGAGAACATC 3477
```

```
Qy 493 MetaArgAlaCysValGlyLeuAlaProGluAsnAsnMet1leLeuGly1leLys 510
Db 3478 ATGAGGCGCTGCTGGTGGCTGGCCCGCAAGAAACATGATCTCGAGTACAAAG 3531

RESULT 5
US-09-677-064-15
/ Sequence 15, Application US/09677064
/ Patent No. 6291224
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Calt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
/ TITLE OF INVENTION: Plants and Uses Thereof
/ FILE REFERENCE: 0706D
/ CURRENT APPLICATION NUMBER: US/09/677, 064
/ CURRENT FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/055, 446
/ PRIOR FILING DATE: 1997-08-11
/ PRIOR APPLICATION NUMBER: 60/055, 526
/ PRIOR FILING DATE: 1997-08-08
/ PRIOR APPLICATION NUMBER: 60/053, 944
/ PRIOR FILING DATE: 1997-07-28
/ PRIOR APPLICATION NUMBER: 09/118, 442
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 15
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-677-064-15

Alignment Scores:
Pred. No.: 2,716-217 Length: 3546
Score: 1813.00 Matches: 443
Percent Similarity: 40.86% Conservative: 31
Best Local Similarity: 38.19% Mismatches: 36
Query Match: 68.88% Indels: 652
DB: 3 Gaps: 9

US-10-718-952-2 (1-510) x US-09-677-064-15 (1-3546)
Qy 1 MetPhe1leGluAsnPheLysValGlyCysProAsnValLys1leThrGluThr1le 20
Db 60 ATGTTTCATCGAGAGCTTCGCGGTGAGAGCCCGCAAGTCCGCTACGGCCGAGATC 119
Qy 21 GlnSerVal1leYrAsn1leGluThr1leThrGluLeuVal1leGluAsnArgAsn1leThr1le 40
Db 120 GAGTGGAGTACCGGTGACACGACGAGAGCTGTGTACGAGGGGCAAGACGCGCTCA 179
Qy 41 GlnThr1leVal1leLysProLysSerVal1leLys1leGluPheLysThrAsn1leHisValPro 60
Db 180 CGTGGTGTCTGCGCCCAAGTCTGTCAAGTACCACTTCGAGCAGAACGCGCTCCC 239
Qy 61 LysLeuGly----- 63
Db 240 AAGCTCGGGTATGTAAGATGACAGGCGCTAGCTCACTCTGTGAACCTCTCTCTC 299
Qy 63 ----- 63
Db 300 CGGTGCTAGTCAATCTCGTGCAGATCAACTGTGCGGTTCCTCTTAATCTTA 359
Qy 63 ----- 63
Db 360 TGAATACTTACTGCTTGTGCTGAAGACGAAACGCTGTAATTGTTGACAGCTACGACAC 419
Qy 63 ----- 63
Db 420 ACTTGCCCATCGGATGCGTCAATCAGCTGATTTGAATTGATTGATGATGCTT 479
Qy 63 ----- 63
```

Db 480 TTCCATATTTCATCATCTTGGCTACTCTGTGCAATGATTACAGAAAGCTCTTTCCTC 539
Qy 63 ----- 63
Db 540 TGAATTTGTCTTAGGCTTTTGTCTGTGACAGTAGCTGATCAATTTGTTCATGTA 599
Qy 63 ----- 63
Db 600 AGATCAAAATTCACAGCAGGACGATGACACAGACAGAACTCATTAACGTAGCAAAATGA 659
Qy 63 ----- 63
Db 660 TACTAGATTACTGCGAAGTGTGCATACGGCGCAATCTGCCATCTGCACCCCTTGT 719
Qy 63 ----- 63
Db 720 TAATTCCTGTCTTANGATGTTGCTACGTGACGCTGTTGTGTATGTGTGACGC 779
Qy 63 ----- 63
Db 780 TGTACCCGCTGTCTCTGTCTGACGAGATGCCAATTTCTGTCTGTGTGTGACAG 839
Qy 64 ValMetLeuValGlyTyrPglYglYAsnAenGlySerThrLeuThrGlyValIleAla 83
Db 840 GTGATGCTGTGTGGGTGGGGAGGCAACACGGGTCCAAGCTGACGGCTGGGTCATTGCC 899
Qy 84 AasnArgGlu----- 86
Db 900 AGCAGGAGATGAGTACTTAATTTGTCTATATTGCTTCCGTGTGTTTCAGTTATTA 959
Qy 87 ----- GlyIleSerThrPalatThr 92
Db 960 ATGGCCTGACAGAACTGAATTTGTGTGGCTGTTCAGGGAGATCTCATGCGCCGACC 1019
Qy 93 LysAspLysIleGlnGlnIleAenYrPheGlySerLeuThrGlnIleSerAlaIleArg 112
Db 1020 AAGCACAAGGTGACAGCAAGCACTACAGGCTC-CTCACCCAGGCTTCCACCATCAAG 1078
Qy 113 ValGlySerPheGlnGlyGlnIleYrYrAlaProPheLysSerLeuPheMet--- 131
Db 1079 GTGCGACGCTACACGGGGAGAGATCTATGCGCGTTCMAAGAGCTCTTCCCATGGTA 1138
Qy 131 ----- 131
Db 1139 ATCTATTATAGACTTGACTAATTACTCTTTTACTGAAACCAACATACATACAAAG 1198
Qy 131 ----- 131
Db 1199 CATATTCGGTAAGTGTAGTTGATGTATTAAGTGAACCTGTCTTTCAGGCCAGTGTTC 1258
Qy 131 ----- 131
Db 1259 TCNAAGTAAACGAATGTATATCATGTGGTTGAAAAAACAAGGTCTAAATTTGTGAAG 1318
Qy 131 ----- 131
Db 1319 GAATGTAAACTTAGCATATAGAAAGGGAGACATGTTAAGAAAGTGTGTAACAAATC 1378
Qy 131 ----- 131
Db 1379 GACTCGGTCTGCCATGTGTGATCTTAATTGCAAGTCAAAAGGTCTGTGGTTAGCTCAA 1438
Qy 131 ----- 131
Db 1439 AGGTTCACAGATTTGATTAACATCTGTGCAGTATTAGAGATGTGTAACGTGTTGA 1498
Qy 131 ----- 131
Db 1499 GATTTCGACATCTCGTGTGTATCTCTTTTCATGACCAAGTGTGAACGTGTTTGA 1558
Qy 132 -ValAenProAspAspIleValPheGlyGlyTyrPheIleSerAasnMetAasnLeuAlaAs 151
|||||

Db 1559 GGTGAACCCAGACGACATTTGTGTGGAGGCTGGGACATTAGCAACATGACCTGGCCGA 1618
Qy 151 PalAmetAlaArgAlaLysValPheAspIleAsnLeuGlnLysGlnLeuArgProTyrIle 171
Db 1619 CTCCATGACACAGGCCCAAGGTCTGTGATTAATTAACCTGCAGAAAGCAAGCTCAGGCCCTACAT 1678
Qy 171 TglIleSerMetLeuProLeu-ProGlyIleTyrAspProAspPheIleAlaAsnGlnG 191
Db 1679 GAGTTCATGTGTGCCACTTCCCGGTATCTATGATCCGACCTTCATCGGGCTAACCAAG 1738
Qy 191 LngIleArgAlaAsnAsnValIleLysGlyThrLysGlnGlnValGlnGlnIleIleL 211
Db 1739 GCTCTGGGCCCAACAGTGTCAATCAAGGCGCACCAAGAAAGAACAGGAGCAGATCATCA 1798
Qy 211 YAspIle----- 213
Db 1799 AGGATATCAGGTATATGATATGATGATGCTAACGTGCTTGTGTCTAAGGTGCAACCAAGT 1858
Qy 213 ----- 213
Db 1859 CAACCTAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1918
Qy 213 ----- 213
Db 1919 ATTTGTTAACGGGTGACCTTAGAAGTCTGCATGTATGTGACATTTGACATTGCGATA 1978
Qy 213 ----- 213
Db 1979 TACAGTACCGCTCACTTGCAATGAGGACTCACAAAAGACTAAACTACTGAAGCTTAA 2038
Qy 213 ----- 213
Db 2039 GCAACTATTCTGATGCTAATGATGATTTGGTGACATGTTGAAGATCTAATTAACGT 2098
Qy 213 ----- 213
Db 2099 GGTGAAGAAATATGCTTCACTAGCACTAAGTAAATTCATACAGAAAGCTATGTA 2158
Qy 214 ----- LysAlaPheLysGlnAlaThrLysValAspLysValValLeuTyrT 230
Db 2159 GCTAAATGAACAGGAGGATTAAGAGAAAGAACAAAGTGACAAAGATGTTGTTGTGTA 2218
Qy 230 hrAlaAsnThrGluArgTyrSerAasnLeuValValGlyLeuAsnAspThrMetGluAsnL 250
Db 2219 CTGCAAAACCTAAGATATAGCAATGTGTGCGCTGTCTCAACGACAGATGGAGAAATC 2278
Qy 250 eulLeuAlaIleValAspArgAsnGlnAlaGlnIleSerProSerThrLeuYrAlaIleA 270
Db 2279 TACTGGCATCTGTGGACAAGACGAGGCGAGATCAACATCAACATATATGCAATTG 2338
Qy 270 IAcYValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValPro 289
Db 2339 CCTGTGTCAATGAGGGGGTGCCTGTCAATCAATGAGGAGCCCCAGAACCTTTGTGCTGT 2398
Qy 289 ----- 289
Db 2399 GTGCGGTGTTGTGTGTGTTGCAAAAGCTTCAAGTGTGCAATTTCTGTCCAAAGTTTC 2458
Qy 290 ----- GlyLeuIle 292
Db 2459 ATGTGTGTATTTCCGTTCCAAAGCTTATTAATCTGTGTGATGTGCTGAGGCTGANT 2518
Qy 293 AspLeuAlaIleAlaArgAsnThrLeuIleGlyIleAspAspPheLysSerGlyGlnThr 312
Db 2519 GATCTTGCTATTAATAAACCACTGCTGATGTGTGTGACGACTTCAAGAGTGGACAGACC 2578
Qy 313 LysMetLysSerValLeuValAspPheLeuValGlyAlaGlyIleLys----- 328
Db 2579 AAGATGAATCTGTCTGTGTGATTTCTTGTGTGTGTGAAT-AAAGGTGGAAACCTTA 2637
Qy 328 ----- 328
Db 2638 GTATCTCTCTTCTATTAAGATGAGATGTTTTTTTGGCAATGACGTTATGCAATTAATCTC 2697

```

QY 329 -----ProThrsSerileValSerTyAsnHisLeuGlyAsn 340
Db 2698 TTCTATATTTTCATTTTCATGCAAGCCCACTCATGCTAGCTACACACACTTGGGAAC 2757
QY 341 AanaapGlyMeCAsnLeuSerAlaprogInThrPheATSerLySGluileSerLySer 360
Db 2758 AACGATGGCATGAACCTGCTCTGCCCTTCAACATTCAAGGATCCCAAGAGGC 2817
QY 361 AenValAlaAspAspMetValAsnSerAsnAlaileuTyGluProGlyGluHisPro 380
Db 2818 AACGTGTGTGATGATGCTGTGAGCAATGCAATCTCTATAGAGCCGCGCAAGCATCC 2877
QY 381 AepHisValValVallelys-----TyrValProTyrVal 387
Db 2878 GATCATGTCGTGTGTATCAAGGTCGTGTAGCTGATCTTCACCTGTTAAAGTTGACAT 2937
QY 388 -----TyrValProTyrVal 392
Db 2938 ATGCAAGGAGATTTACATTGAACCTGTCACTCTTTGTCAGATATGTCGGTACGTG 2997
QY 393 GLVAspSerLyAspAlaMetAspGluTyThrSerGluilePheMetGlyGlyLySer 412
Db 2998 GGAAGACAGCAAGAGGCTATGACAGTACCTCAGAGATCTTCATGGCGCAAGAAC 3057
QY 413 ThrileValleuHisAsnThrcysGluAspSerleuValAlaProIleileuAsp 432
Db 3058 ACCATGTCCTGCACAAACCTGTGAGGACTCGCTCTCGCGCACTTCACTCTGAT 3117
QY 433 LeuValleuValleuAlaGluLeuSerThraGllleGluPheValleuGluAnglu 450
Db 3118 CTGGTCTCTTGTGCTGAGCTCAGACACAGATCAGCTGAAAGCTGAGGAGAGGTAA 3177
QY 450 ----- 450
Db 3178 GCCCCCCAAGTATTAACCTGAAGCAACCTGCAGCTGAGTATAGCACTTTAATA 3237
QY 451 -----GlyLyPheHisSerPheHisProValAlaThrile 462
Db 3238 CCTTCGTGTCTCTCTTATGAGAGCAATTCACCTCTTCACCCGGTGGCAACATC 3297
QY 463 LeuSerTyLeuThryleuAlaProleu----- 471
Db 3298 CTGAGCTACCTACCAAGGACACCTCGTAAAGCTTTCTCTGCATCCCGCATCACTG 3357
QY 472 -----Val 472
Db 3358 CACTGCGTTTGTTCATTCACGCACTGATGCTCTCTCGAAACCTGAACAACAGGT 3417
QY 473 ProProGlyThrProValAlaAsnAlaLeuSerLySGluAlaMetLeuGluAsnile 492
Db 3418 CCCCCTGGCAACCGGTGTGAACGCTCTGGCCAAAGCAACGCGCATGTGGGAACATC 3477
QY 493 MetAlaGlaCysValGlyleuAlaProGluAsnAsnMetileuGluTyTylys 510
Db 3478 ATGAGGCGCTGTGGTGGCTGGCCCAAGAAACATGATCTCTGAGTAAAG 3531

```

RESULT 6
US-09-118-442-14

```

; Sequence 14, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Wang, Hongyu
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446

```

```

; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
; US-09-118-442-14

Alignment Scores:
Pred. No.: 3,62e-217 Length: 3546
Score: 1812.00 Matches: 442
Percent Similarity: 40.86% Conservative: 32
Best Local Similarity: 38.10% Mismatches: 36
Query Match: 68.84% Indels: 652
DB: 3 Gaps: 9

US-10-718-952-2 (1-510) x US-09-118-442-14 (1-3546)
QY 1 MetPheileGluAsnPhelyValGluCyProAsnVallyTyThrGluThrGluile 20
Db 60 ATGTCATCGAGAGCTTCGCGTCGAGAGCCGCCAGCTGAGTACGCCGAGAGATC 119
QY 21 GlnSerVallyTyAsnTyGluThrThrgluValleuValleuHisGluAsnArgAsnGlyTyThr 40
Db 120 GAGTCGAGTACCGGTATGACACAGAGAGCTGTACAGAGGAGCAAGAGAGCGGCCCTCA 179
QY 41 GlnTrileValleuAlaProLySerVallyTyGluPheLyThraHisValleuPro 60
Db 180 CGCTGGTGTCTCGCCGCCAAGTCCGTCAAGTACAACTCCGAGCAGAAACCGCGTCCCC 239
QY 61 LysleuGly----- 63
Db 240 AAGCTCGGGTATGTACGAGATGACGAGCGCCCTAGCCTCACTCTGTGAACCTCTCTC 299
QY 63 ----- 63
Db 300 CCGTCTCACTCAATCTCTCCGTGAGATCAACTGTGCGGCTTCTCTTAATCTTA 359
QY 63 ----- 63
Db 360 TGAATATCTTACTGCTTTGCTGAAGACGAACCGTGAATTTGTGACAGTACGACAC 419
QY 63 ----- 63
Db 420 ACTGCCCATCGAGTGGCTCAATCACTGATGATTTGAATTCGATGATGTCCTT 479
QY 63 ----- 63
Db 480 TTCCATATTTGCATATCCCTCGCTACTGTGCAATGATTAACAGAAAGTCTTTCTC 539
QY 63 ----- 63
Db 540 TGAATTTGTCTTAGCTTTTGTCTGTGACGAGAGCTGGIATCAATTTGTCATGTA 599
QY 63 ----- 63
Db 600 AGATCAATTCGACGAGGAGAGATGAGACAGACAGAACTCATTTACCTAGCAATTTGA 659
QY 63 ----- 63
Db 660 TACTAGATTACTGCAAGTGTGATACGGCGCAATTCGCATCTGAGACCCCTTTGT 719
QY 63 ----- 63
Db 720 TAATTCCTGTCATGATGATGTGCTTACGTGACAGCTGTTGTGTATGTGTGACAGC 779
QY 63 ----- 63

```

Dh 780 TGTACGCCCTTGTCTCTGTCCGACGAGANTGCCAACTTTTCTGTCTGTGTGTGACAG 839
Qy 64 ValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyValIleAla 83
Db 840 GTGATGCTTGTGGGTGGGAGGCAACAAGGGTCCACGTCAGCGGTGGGTCAATGGCC 899
Qy 84 AsnArgGlu----- 86
Db 900 AACAGGAGTGAAGTACTTAATTGTCTAATTGCTTCCGTTGTTTTCAGTTATTA 959
Qy 87 -----GlyIleSerTyrAlaThr 92
Db 960 ATGGCTTAACAGAACTGAATTTGTGTGTGTGTTCAGGGGATCTCATGGCCGACC 1019
Qy 93 LysAspLysIleGlnGlnAlaSerTyrPheGlySerLeuThrGlnAlaSerAlaIleArg 112
Db 1020 AAGGACAAAGGTGCAGACAACTACTACGGCTC-CTCACCCAGGCTTCCACCATCAGA 1078
Qy 113 ValGlySerPheGlnGlyGluGluIleTyrAlaProPheLysSerLeuPheProMet---- 131
Db 1079 GTCCGACGCTACACAGGGAGAGATCTATGCCCGTTCAAGAGCCTCTCCCATGGTA 1138
Qy 131 ----- 131
Db 1139 ATCTATTATAGACTTGACTAATACCTCTCTTTTACTGAACCAACATACATACAAAG 1198
Qy 131 ----- 131
Db 1199 CATATTCGGTAAGGTGTAGTTGANTGTATAAATGAACCTGTCTTTCAGGCCAGTGTTC 1258
Qy 131 ----- 131
Db 1259 TCAGATGAACGGAATCTAATCATTTGGGTGAAAAAACAAGGTCTAATTTTGTGAAG 1318
Qy 131 ----- 131
Db 1319 GAAAGTTAACTTAGCATATGATAAGGAGGAGACACTGTAAAGAAAGGTGTGAAACATC 1378
Qy 131 ----- 131
Db 1379 GACTCGGTCTGCATGTGTGTGATCTCTAATTGCAGAGTCMAAAGTCTGTGTAGCCAA 1438
Qy 131 ----- 131
Db 1439 AGGTTCACAGACTTTGGATTACACTGTGCAGATTAAGCATGTGTCTAACTGGTTGCA 1498
Qy 131 ----- 131
Db 1499 GATTGACAGACTCGGTGTGTGTATCTTTTCATGACCAAGTGTAAACGTGTTTCA 1558
Qy 132 -ValAsnProAspAspIleValPheGlyGlyTyrAspIleSerAsnMetAsnLeuAlaAs 151
Db 1559 GGTGAACCCAGACGACATGTGTTCGAGAGCTGGGACATTAGCAACATGAACTGGCCGA 1618
Qy 151 PAlaMetAlaArgAlaLysValPheAspIleAspLeuGlnLysGlnLeuArgProTyrPhe 171
Db 1619 CTCCATGACCAAGGCTCAAGCTGTGATATTAACCTGCAGAGCAGCTCAAGCCCTACAT 1678
Qy 171 GlnSerMetLeuProLeu-ProGlyIleTyrAspProAspPheIleAlaIleAsnGlnG 191
Db 1679 GGAAGTCATGTGGCCACTTCCCGGTATCTATGATCCGACCTTCACTGGGCTTAACGAG 1738
Qy 191 LngLysAlaArgAlaAsnAsnValIleLysGlyTyrLysGlnGlnGlnValGlnGlnIleLe 211
Db 1739 GCTTCGGCCCAACAGTGCATCAAGGCGCAACAAGAAAGAACAGGTGAGCAGATCATCA 1798
Qy 211 LysAspIle----- 213
Db 1799 AGGATATCAGGTATATGATATGATGCTTAACGTGCTTGGTCTAAGGTGACCCAGTGG 1858
Qy 213 ----- 213
Db 1859 CAACCTAAAAAATAATAATACTATGAAATTTGGTAAATATATACATATCAGAGCAT 1918

Qy 213 ----- 213
Db 1919 ATTTGTTAAACGGTGCATTTAGAGAGTCTGCATGTATGTTGGACAATTTGACATTCGATA 1978
Qy 213 ----- 213
Db 1979 TACAGTACCGCTCACTTCGATGAGAGCTCCACAAAGAACTAAACTACTGAAGACTTAA 2038
Qy 213 ----- 213
Db 2039 GCAACTATTCGTAGCTAATGATGATATTGTGTGACATGTGTTGAAGATCTAGATTACGT 2098
Qy 213 ----- 213
Db 2099 GGTGAAGAAATATAGTTCTACTAGTATTAATGATTCATTACAGAACATGGCTTATGTA 2158
Qy 214 -----LysAlaPheLysGluAlaThrLysValAspLysValValIleLeuTyrTrp 230
Db 2159 GCTAATGAAACAGGAGTTTAGGAGAGAGAACAAAGTGAACAGATAGTTGTGTGTGGA 2218
Qy 230 hrAlaAsnThrGluArgTyrSerAsnLeuValValGlyLeuAsnAspThrMetGluAsnL 250
Db 2219 CTGCAAAACACTGAAGGTATAGCAATGTGTGGCTGTCTCAACGACACAGATGAGAAATC 2278
Qy 250 euLeuAlaAlaValAspArgAsnGluAlaGluIleSerProSerThrLeuTyrAlaIleA 270
Db 2279 TACTGGCATCTGTGACAAAGACGAGCGAGGTATCACCATTCACATATATGTCATTGG 2338
Qy 270 LAcYValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValPro- 289
Db 2339 CCTGTCTATGAGAGGGGGTCCGTTTCATCAATGGAGCCCAAGAACACTTTGTGCTTG 2398
Qy 289 ----- 289
Db 2399 GTGCGTGTGTGTGTGTGTGCAAAAAGCCTCATGGNGTTGCATTTCTGTCCAAAGTTTC 2458
Qy 290 -----GlyLeuIle 292
Db 2459 ATGGTGTGTATTTCTGTCCAGGCTTATTAATACCTGTGTGATGTTCGTAGGGCTGATTT 2518
Qy 299 AspLeuAlaIleAlaArgAsnThrLeuIleGlyGlyAspAspPheLysSerGlyGlnThr 312
Db 2519 GATCTTGCTATAAAAAACAACACTGCTTGATGTGGTGGAGACATTCAGATGGAGACACC 2578
Qy 313 LysMetLysSerValLeuValAspPheLeuValGlyAlaGlyIleLys----- 328
Db 2579 AAGATGAATCTGTGTGTGTGATTTCTGTGTGTGTGTGAAT-AAAGGTGGAACCTA 2637
Qy 328 ----- 328
Db 2638 GTATCTCTTCTATTAAGATGAAGTGTGTTTGTGCAAAATGACCTTATTCGAATTAATC 2697
Qy 329 -----ProThrSerIleValSerTyrAsnHisIleuGlyAsn 340
Db 2698 TTCTATATTTTCATTTTCATGACGCCACCTCAATCGTAGCTACCAACACTTGGGAAC 2757
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyIleSerLysSer 360
Db 2758 AACGATGGCATGAACCTGTCTGCCCTTCAACATTCAGGTCCAAAGAGATCTCCAAAGAC 2817
Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlyLysPro 380
Db 2818 AACGTGTGTGTGATGATGTGTGTGAGCAATGCAATCTCTATGAGCCCGGCGAGATCC 2877
Qy 381 AspHisValValValIleLys----- 387
Db 2878 GATCATGTGTGTGTGTATCAAGGTCTGTTAAGCTGATCTTTCACTGTGTTAAAGTTGACAT 2937
Qy 388 -----TyrValProTyrVal 392
Db 2938 ATGCAAGGACATTTACATTTGAACCTGTGCACTCTTTTGTGACGATATGTGCGTACGTG 2997

QY 393 G1yAspSerLySaRgAlaMetAspG1uTyRThrSerGluLePheMetG1yG1yLysSer 412
DB 2998 GGAGACAGCAAGAGGCGTATGACGATCACCTCAGAGATCTTCATGGCGCAAGAAC 3057
QY 413 ThrLeValLeuH1eAsnThrCySeG1uAspSerLeuLeuAlaAlaProIleIleLeuAsp 432
DB 3058 ACCATCGTGTCTCACAAACCTGTGAGACTGTCTCGCCGACATCATCTTGAT 3117
QY 433 LeuValLeuLeuAlaG1uLeuSerThrArgIleGluPheLySa1eGluAsn----- 449
DB 3118 CTGGTCTCTTGGCTGAGTCAACAGATCCAGCTAAAGCTGAGGAGGGGTAGA 3177
QY 449 ----- 449
DB 3178 GCGCCCAAGTATTAACCTGAAGACGCTGACGCTAGTGAATACACTTTAATA 3237
QY 450 -----GluG1yLysPheH1eSerPheH1eProValAlaThrIle 462
DB 3238 CCTTGTGTGTCTCTTATGAGAGCAAAATTCACCTCTCCACCCGGTGGCCACATC 3297
QY 463 LeuSerTyRLeuThrLySa1aProLeu----- 471
DB 3298 CTGAGCTACTCAACAGGACCCCTGTAAAGCTTTCTCTGCAATCCCGCATCAGT 3357
QY 472 -----Val 472
DB 3358 CACTGGCTTTGCTTCAATCCAGCCACTGATGCTCTCTTGAACCTGAACAGAGTT 3417
QY 473 ProProG1yThrProValValaMetLeuSerLySa1eArgAlaMetLeuLeuMetIle 492
DB 3418 CCCCCGAGACACCGGTGTGAACGCTGCGCCAAACAGACGCGCATGTGGAAGAACATC 3477
QY 493 MetArgAlaCySaValG1yLeuAlaProG1uAsnAsnMetIleLeuG1uTyR1yS 510
DB 3478 ATGAGGCGCTGCGTGGCTGGCGCCCAAGAACATCATCTCTGAGATCAAG 3531

RESULT 7
US-09-677-064-14
; Sequence 14, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Calc, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-14

Alignment Scores:
Pred. No.: 3,62e-217 Length: 3546
Score: 1812.00 Matches: 442
Percent Similarity: 40.86% Conservative: 32
Best Local Similarity: 38.10% Mismatches: 36
Query Match: 68.84% Indels: 652
DB: 3 Gaps: 9

US-10-718-952-2 (1-510) x US-09-677-064-14 (1-3546)
QY 1 MetPheIleGluAsnPhelySaValG1uCySProAsnValLyTyRThrGluThrGluIle 20
DB 60 ATGTTTCAATCGAGAGCTTCGGGTGAGAGCCGCCACGTCGGGTACAGGCCGAGATTC 119
QY 21 GlnSerValTyRAsnTyRGluthrThrGluLeuValH1eGluAsnArgAsnG1yThrTyR 40
DB 120 GAGTCGAGATCCGGGTACAGACGACGAGCTGTACAGAGGAGAGAGACGCGCTCA 179
QY 41 GlnTPH1eValLySaProLySaValLyTyRThrGluPheLySaH1eH1eValPro 60
DB 180 CCGTGGGTGTCGCCCAAGTCGTCAAGTACAACTTCGAGACCAAGCCGCTCC 239
QY 61 LysLeuGly----- 63
DB 240 AAGCTCGGGTATGTAACGATGACGGGCGCTAGCTCACTCTGTGAACCTCTCTCTC 299
QY 63 ----- 63
DB 300 CCGTCTCAGTCAAACTCTCCGTCGAGATCACTGTGCGGCTTCTCTTAATCTTAA 359
QY 63 ----- 63
DB 360 TGAATACTTACTGCTTGGCTGAAAGACGAACCGTGTATTTGTGACAGCTACGACAC 419
QY 63 ----- 63
DB 420 ACTTGCCATCCGGATGGGTCAATCAGCTGATTTGAATTCATCGATGATGTCCTT 479
QY 63 ----- 63
DB 480 TTCATATTTGCATCATCTCTGCTGATGTCATGATTAACAGAAAGTCTTTCTCTC 539
QY 63 ----- 63
DB 540 TGAATTTGTCTTAAAGCTTTTGTCTGTGACGTGAGCTGTATCAATTTGTATGTA 599
QY 63 ----- 63
DB 600 AGATCAATTCACAGAGGAGATGACGACGACAGACAACTATTACACTAGCAATTTGA 659
QY 63 ----- 63
DB 660 TACTAGATTAATGCGCAAGTGTGATAGCGCGCAATTCCTGACCCCTTTGTT 719
QY 63 ----- 63
DB 720 TAATTCCTGTTCTATGCAATGTTGCTTACGTGACGCTGTTGTATGATGTCAGGC 779
QY 63 ----- 63
DB 780 TGTACGCGCTTGTCTCTGTCCGACGATGATGCCAATTTCTGTCTGTGTGTCAGC 839
QY 64 ValMetLeuValG1yTyRGlYAsnAsnG1ySerThrLeuThrG1yG1yValIleAla 83
DB 840 GTGATGCTTGTGGGTGTGGGAGGCAACACGGGTCCAGCTGACGCTGGGTGTATGTC 899
QY 84 AsnArgGlu----- 86
DB 900 AACAGGAGTGAATGATTAATTTGCTTATTTGCTTTCCTTTCCTTTCAGTTATTA 959
QY 87 -----GlyLysSerTrpAlaThr 92
DB 960 ATGCGCTAACAGAACTGAATTTTGTGTGTGTTTTCAGGGATCTCATGGCGAC 1019
QY 93 LysAspLySa1eGlnAlaAsnTyRPhelG1ySerLeuThrGlnAlaSerAlaIleArg 112
DB 1020 AAGGACAGAGTGCACAGCAAGCACTACGAGCTC-CTCACCCAGGCGCTCCACATCAGA 1078
QY 113 ValG1ySerPheGlnG1yGluGluIleTyRAlaProPheLySaSerLeuAsnProMet--- 131
DB 1079 GTTCGAGCTCAACGGGAGGAGATCTATGCGCGCTTCAAGAGCTCTCTCCATGTTA 1138

QY 131 ----- 131
Db 1139 AACTATTATAGACTTGACTTAATCTCTTTTCTTACTGAAACCAACATACATACAAAG 1198
QY 131 ----- 131
Db 1199 CATATTCCGTAAGTGTCTACTGTATGTATAAATGAACCTGTCTTTACAGCCAGTGTTC 1258
QY 131 ----- 131
Db 1259 TCAGATAAAGCAATGTATATCATTTGGGTGAAAAAACAAGGCTTAATTTTGTGAAG 1318
QY 131 ----- 131
Db 1319 GAAAGTTAAACCTTAGATAATGAAGGGAGACCTTAGAAGAGTGTGAACAATC 1378
QY 131 ----- 131
Db 1379 GACTGGGTCTGCATGTGTGATCTTACTTGCAAGTCAAAAAGTCTGTGTGTTAGCCAA 1438
QY 131 ----- 131
Db 1439 AGGTTCAGCATCTTTGATTAACACTGTGAGATTGACATGATGAGATGCTAACTGTGCA 1498
QY 131 ----- 131
Db 1499 GATTGCGAGACTGGTGTGTGTATCTTCTTTTCATGACCAAGTGTAACTGGTTTTC 1558
QY 132 -ValaenProaspAsp11eValPheglYglYTrpaSp11eSerAsnMetAsnLeu11aAs 151
Db 1559 GGTGAACCCAGACGACATTTGTTCGAGGCTGGGACATTTAGACATGAACTGGCCGA 1618
QY 151 pAlaMetC1aArgAlaAsnVal11eLeuValPheAsp11eAspLeuGlnLysGlnLeuArgProTyr 171
Db 1619 CTCATGACACAGGCGCAAGGTGCTGATATGACCTGACAGAACAGTCAAGCCCTACAT 1678
QY 171 fcGusMetLeuProLeu-ProGly11eTyrAspProaspPhe11eAla1aAsnGlnG 191
Db 1679 GAGGTCCATGTGTCCACTTCCCGGTATCTATGATCCGACTTCATCGCGCTAACCAAG 1738
QY 191 Juc1uArgAlaAsnVal11eLeuGlyThrLysGlnGlnValGlnGln11eLeu 211
Db 1739 GCTCTGCGCCACAGTGTCAATCAAGGGCACAGAAAGACAGGTGAGACAGATCTCA 1798
QY 211 yAsp11e----- 213
Db 1799 AGGATATCAAGTATATGATGATGATGCTTAACGTGCTGTCTAAAGTGACCCAGTG 1858
QY 213 ----- 213
Db 1859 CAACCTMAAACAAATTAATACTACTATGATTTGTAAATATACATACATACAGACAT 1918
QY 213 ----- 213
Db 1919 ATTGTTTAAACGGGTGACTTAGAGTCTGCATGATGTTGACAAATTTGACATTCGATA 1978
QY 213 ----- 213
Db 1979 TACAGTACCGGCTCACTTGATGAGAGCTCCACAAAGAACTAAACTAGTGAAGCTTAA 2038
QY 213 ----- 213
Db 2039 GCAACTATTCTAGTATATGATGATTTGTGTGACATGTTGAAAGTCAATTAACGT 2098
QY 213 ----- 213
Db 2099 GGTGAGAAATATGTGCTACATAGTATTAAGTAATTCATACAGAACAAATGCTTATGTA 2158
QY 214 -----LysAlaPheLysGlnAlaThrLysValAspLysVal11eLeuTyrP 230
Db 2159 GCATATGAACACGGAGTTTAGGGAGAGAAGAACAAAGTGGACAAAGATAGTTGTGTGGA 2218

QY 230 hrAlaAsnThrGluArgTyrSerAsnLeuVal11eGlyLeuAsnAspThrMetGluAsnL 250
Db 2219 CTGCACACACTGAAGAGATATAGCAATGTGTGCGCTGTCTCAACGACAGATGAGAAATC 2278
QY 250 euLeuAlaAlaValAspArgAsnGluAlaGln11eSerProSerThrLeuTyrAla11eA 270
Db 2279 TACTGCATCTGTGGACAAAGAACAGAGCGGAGGTATCACATCAACATATATAGCATTTG 2338
QY 270 JAcysValMetGluAsnValProPhe11eAsnGlySerProGlnAsnThrPheValPro 289
Db 2339 CCTGTGTATGAGGGGGGTGCGCTTCATCAATGGAGGCCCCAGAACACCTTTGTCTCTG 2398
QY 289 ----- 289
Db 2399 GTGCGTGTGTGTGTGTGTGCAAAAGCCTCATGTGTTCATTTCTGTCCAAAGTTTC 2458
QY 290 -----GlyLeu11e 292
Db 2459 ATGTGTGTATTTCTGTCCAAAGCTTATTAACCTGTGCAATGTCGTAGGGCTGATT 2518
QY 293 AspLeuAla11eAlaArgAsnThrLeu11eGlyLysAspAspPheLysSerGlyInthr 312
Db 2519 GATCTTGATATAAAAAACAAGCTTGATTTGGTGTGACGACTTCAAGATGAGACAGACC 2578
QY 313 LysMetLysSerValLeuValAspPheLeuValGlyAlaGly11eLys----- 328
Db 2579 AAGATGAATCTGTCTGTGTGATTCATTTCTTGTGTGTGTGAAT-AAAGTGGAACTTA 2637
QY 328 ----- 328
Db 2638 GATATCTCTTCAATTAAGTAAAGTCTTTTGGCAAAATGACGTATTTGCAATAATC 2697
QY 329 -----ProThrSer11eValSerTyrAsnH1sLeuGlyAsn 340
Db 2698 TTCTATATTTCAATTTTCAATGACGACCCACCTCAATGTAGACTCAACACCACTTGGGAAAC 2757
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnInthrPheArgSerLysGln11eSerLysSer 360
Db 2758 AACGATGGCATGAACCTGTGTGCTTCAAACTTCAGGTTCAAAGAGATCTTCAAGAGC 2817
QY 361 AsnValValAspAspMetValAsnSerAsnAla11eLeuTyrGlnProGlyGlnH1sPro 380
Db 2818 AACGTGTGTGATGACATGTTCTGTGAGCAATGCCATCTCTATGAGCCCGGACAGATCCC 2877
QY 381 AspHisValVal11eLys----- 387
Db 2878 GATCATGTCTGTGTATCAAGGCTGTGTAGTGAATCTTCACTCGTAAAGTTGACAT 2937
QY 388 -----TyrValProTyrVal 392
Db 2938 ATGCAAGGAGATTATCATTTGAAACTTGTCACTCTTTTGTGCAGATATGTGCCGTAAGTG 2997
QY 393 GlyAspSerLysArgAlaMetAspGluTyrThrSerGln11ePheMetGlyLysSer 412
Db 2998 GAGACACGACAAAGGCTATGACAGATACACTCAGAAATCTTCAATGAGCGGCAAGAAC 3057
QY 413 Thr11eValLeuHisAsnThrCysGluAspSerLeuLeuAla1aPro11eLeuAsp 432
Db 3058 ACCATGTGCTGCACAAACACTGTGAGACTGCTCTCGCGCACTATCATCTTGAAT 3117
QY 433 LeuValLeuLeuAlaGluLeuSerThrArg11eGluPheLysAlaGluAsn----- 449
Db 3118 CTGTGTCTTGTGTGAGCTGACGACCAAGATTCAGCTAAAGCTGAGGAGGGGTAGA 3177
QY 449 ----- 449
Db 3178 GCCCCCAAGTATTAACCTGAAGACGCTGACGCTAGGTATAGCACTTTTAATA 3237
QY 450 -----GluGlyLysPheH1sSerPheH1sProVal1aThr11e 462
Db 3238 CCTTGTGTGTCTCTTATGAGGACAAATTCACCTCTTCAACCGGTGCGCACATC 3297
QY 463 LeuSerTyrLeuThrLysAlaProLeu----- 471

```
Db      3298 CTGAGCTACTCCACCAAGGACCCCTGGTAAGCTTTCTCTCGATCCCGGCATCACTG 3357
Qy      472 -----Val 472
Db      3358 CACTGCGTTTCTTCATCCACCACTGATGCTCTCTTGAACCTGACACAGGTT 3417
Qy      473 ProProGlyThrProValValAsnAlaLeuSerLysGlnArgAlaMetLeuGluAsnIle 492
Db      3418 CCCCCTGGACACCGGTGGTGAACGCTCTGGCCAGACAGCGCGATGCTGGAGAACATC 3477
Qy      493 MetAlaGlaCysValGlyLeuAlaProGluAsnAsnMetIleLeuGluTyrIlys 510
Db      3478 ATGAGGCGCTGCGTTGGCTGGCCCGACAGAACCAATGATCTCGAGTAAAG 3531

RESULT 8
US-09-734-237B-72
/ Sequence 72, Application US/09734237B
/ Patent No. 6818752
/ GENERAL INFORMATION:
/ APPLICANT: Rozzell, J. David
/ APPLICANT: Bui, Peter
/ APPLICANT: Hua, Ling
/ TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
/ FILE REFERENCE: B583:40608
/ CURRENT APPLICATION NUMBER: US/09/734,237B
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 09/494,921
/ NUMBER OF SEQ ID NOS: 79
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 72
/ LENGTH: 1602
/ TYPE: DNA
/ ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-72

Alignment Scores:
Pred. No.: 1,31e-163 Length: 1602
Score: 1382.50 Matches: 277
Percent Similarity: 69.27% Conservative: 86
Best Local Similarity: 52.86% Mismatches: 140
Query Match: 52.53% Indels: 21
Gaps: 8

US-10-718-952-2 (1-510) x US-09-734-237B-72 (1-1602)
Qy      3 IlegluAsnPhelYValGluCysProAsnValIysTyrThrGluThrGluIleGlnSer 22
Db      25 ATCACCCTCCGTTAAAGTACTTACCGACAGTGCACGTCACAGCAACGAGCTGCTCAC 84
Qy      23 ValTyrAsnTyrGluThrThrGluLeuValIleGluAsnArgAsnGlyThrTyrGlnTyr 42
Db      85 AAGTACAGCTACGAAATATCT---GTAGTTACGAAACACGTAAGCGCGCTTCGAT--- 138
Qy      43 ILeuValIleProLysSerValIysTyrGluProLysTyrAsnIleHisValPro---Lys 61
Db      139 ---GTAGACCCCACTGTCACAGCTACGCTTCAAACTGACTGAAAACCGCGGAAAA 195
Qy      62 LeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyVal 81
Db      196 CTAGGAATTAATGCTCTTGGGTGAGGTGCAACAAAGGCTCCACTTATGCGCTCGGTA 255
Qy      82 ILeuAlaAsnArgGluGlyIleSerTProAlaThrIleAspLysIleGlnGlnAlaSerTyr 101
Db      256 TTGGGCAATATGACACATGTGGAGTTTCAAACTTAAAGGAGCGGTTGACAACTAC 315
Qy      102 PheGlySerLeuThrGlnAlaSerAlaIleArgValGly---SerPheGlnGlyGlnGlu 120
Db      316 TTCGGCTCATGACTCAATGCTTCTACCTTGAAGAACTGGTATGATCGGAGGGAATGAC 375
Qy      121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspIleValPheGly 140
Db      121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspIleValPheGly 140
```

```
Db      376 GTTATGCTCTCTTTAACTCTCTGTGGCCATGCTTACCCCAACGACTTTGTGCTCT 435
Qy      141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValPheAsp 160
Db      436 GGTTCGACATCATTAACGACAGATCTATACGAGTATACGAGAGAAAGTCTTCGAA 495
Qy      161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
Db      496 TATGATCTGCAACACGCTTGAAGCGCAAGATGCTCTTGGAGGCCCTTCTTCAT 555
Qy      181 TyrAspProAspPheIleAlaIleAsnGlnGluIleValArgAlaAsnAsnValIle----- 198
Db      556 TACTACCTGATTTATTTAGCTATATCAAGATGAGAGCCATATACGATCAATTTG 615
Qy      199 -----LysGlyThr-----LysGlnGlnValGlnGlnIleIleLys 211
Db      616 GATGAAAAAGGACGTAACCAACGAGGGGTAAAGTGAACCATCTGCACGATCAGACGC 675
Qy      212 AspIleLysAlaPheLysGluAlaThrLysValAspLysValValIleLeuThrAla 231
Db      676 GATATCCAGAAATTCAAAGAAAGAAACCGCTTGATTAAGTATCGTTCTTGGACTGCA 735
Qy      232 AsnThrGluArgTyrSerAsnLeuValGlyLeuAsnAspThrMetGluAsnLeuLeu 251
Db      736 AATACTGAGAGGTACGTAAAGTATCTCTGCTGTATTAAGACCATGAAACCTCTTG 795
Qy      252 AlaAlaValAspArgAsnGluAlaGluIleSerProSerThrLeuTyrAlaIleAlaCys 271
Db      796 CAGTCTATTAAGAAAGACATGAAGAGATGCTCTTCCAGATCTTTCAGACACACATCT 855
Qy      272 ValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValProGlyLeu 291
Db      856 ATCTTGAAAGGTGTCCTTATATATGTTTCAACGCAAGTACTTGTGTTCCCGCTTG 915
Qy      292 IleAspLeuAlaIleAlaArgAsnThrLeuIleGlyIleAspAspPheLysSerGlyGln 311
Db      916 GTTCAGCTGCTGACACAGAGGATACATTCATTGCGGAGACATCTCAATGCGGACAA 975
Qy      312 ThrLysMetLysSerValLeuValAspPheLeuValGlyAlaGlyIleLysProThrSer 331
Db      976 ACCAAGTTGAAGTCTGTTCTGGCCAGTCTTATGAGATGACGATTAACCGGCTTCC 1035
Qy      332 IleValSerTyrAsnHisLeuGlyIleAsnAsnAspGlyMetAsnLeuSerAlaProGlnThr 351
Db      1036 ATTGATCTCTTAACCATTTAGGCAATATGACGCTTAACTTATCTGCTCCAAACAA 1095
Qy      352 PheArgSerLysGluIleSerLysSerAsnValAlaAspAspMetValAsnSerAsnAla 371
Db      1096 TTATAGCTTAAGAGATTTCCAAAGTTCTGTGTCATAGATGACATCATCGCTTAATGAT 1155
Qy      372 IleLeuTyrGluPro-----GlyGluHisProAspHisValValIleLysTyrVal 389
Db      1156 ATCTTGTCATATGATTAACCTGGGTAAAGGTTGACCACTGATGTCATCAATATATG 1215
Qy      390 ProTyrValGlyAspSerLysArgAlaMetAspGlyTyrThrSerGluIlePheMetGly 409
Db      1216 AAGCCGTCGGGAGCTCAAAAGTGCACATGACGAGTATTCAGAGAGTGTATGTTAGT 1275
Qy      410 GlyLysSerThrIleValLeuHisAsnThrCysGluAspSerLeuLeuAlaIleProIle 429
Db      1276 GGCCATTAACCGGATTTCCATTTACATGTTTGCAGAAATTTCTTATCTGCTACGCGCTTG 1335
Qy      430 IleLeuAspLeuValLeuLeuAlaGluLeuSerThrArgIleGluPheLys----- 446
Db      1336 ATCATCGATCTTTAGTACATGATGAGTTTGTACAAGATGCTCTATTAAGAAAGTGCAC 1395
Qy      447 -----AlaGluAsnGluGlyLysPheHisSerPheHisProValAlaThrIleLeuSer 464
Db      1396 CCAGTTAAAGAAAGATGCGCAAAATTCAGAACTTTTATCCAGTTTAACTTCTTGAGT 1455
Qy      465 TyrLeuThrLysValProLeuValProProGlyThrProValValAsnAlaLeuSerLys 484
Db      1456 TACTGGTTAAAGCTCCATTAACAAAGACCAAGATTTCAACCGGATGAGGCTTAAACAG 1515
```


QY 485 G1AARGAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaProGluAsnAsn 504
DB 1516 CAAGAAGCCGCTTGAATAATTTTAAAGATTGTTGATTGATGCTTCTTCAAAACGAA 1575
QY 505 MetIleLeuGlu 508
DB 1576 CTAAGATTGAA 1587

RESULT 9
US-09-734-237B-74
Sequence 74, Application US/09734237B
Patent No. 6818752
GENERAL INFORMATION:
APPLICANT: Rozzell, J. David
APPLICANT: Bul, Peter
APPLICANT: Hua, Ling
TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
FILE REFERENCE: B583:40608
CURRENT APPLICATION NUMBER: US/09/734,237B
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 09/494,921
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn version 3.1
SEQ ID NO 74
LENGTH: 1605
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic gene derived from Saccharomyces cerevisiae myo-inositol
OTHER INFORMATION: 8 encoding the same amino acids to reduce free energy of folding,
OTHER INFORMATION: and a gly codon inserted after the initiating met codon
US-09-734-237B-74

Alignment Scores:
Pred. No.: 1,32e-163 Length: 1605
Score: 1382.50 Matches: 277
Percent Similarity: 69.27% Conservative: 86
Best Local Similarity: 52.86% Mismatches: 140
Query Match: 52.53% Indels: 21
DB: 4 Gaps: 8

US-10-718-952-2 (1-510) x US-09-734-237B-74 (1-1605)

QY 3 I1eGluAsnPhelYeuValGluCysProAsnValIysThrGluThrGluIleGlnSer 22
DB 28 ATCACTCTCGTTAAAGTTGTAAGTGAACAATGACTTACAAAGACACACAGCACTGCTGACT 87
QY 23 ValIYrAsnTYrGluThrThrGluLeuValI1eGluAsnArgAsnGlyThrTYrGluInTrp 42
DB 88 AAATCTCTTACGAAACCGCT---GTAATTAATAAAAGCTTCTGCTGCTTCGAT--- 141
QY 43 I1eValYsProLysSerValIysTYrGluPhelYeuThrAsnIleHisValPro---Lys 61
DB 142 ---GTAATCTCAAGCTTCAAGACTGATTAATCAACCTGATCAAGAAACCGGAAAG 198
QY 62 LeuGlyValMetLeuValGlyTYrGlyIYrAsnAsnGlySerThrIleuThrGlyVal 81
DB 199 CTGGGATCATCTGATCGGCTCGGCTGCTGTAACAAGCGCTCTACTCTGTTGCTGATCTGTT 258
QY 82 I1eAlaAsnArgGluGlyI1eSerTrpAlaThrIYrAsnLysI1eGlnGlnI1aAsnTYr 101
DB 259 CTGGCAACCAACACACACATTAATTCAGACTAAGAGAGTGTAAACAGCCGAACTAC 318
QY 102 PheGlySerLeuThrGlnI1aSerAlaI1eArgValGly---SerPheGlnI1eGlyGlu 120
DB 319 TTGGTCTATATAGCTAGGTTCTACTCTGAAGCTGGGCAATTGATGCTGAAGGTAAACGAC 378
QY 121 I1eTYrAlaProPheLysSerLeuLeuProMetValAsnProAspAspI1eValPheGly 140
DB 379 GTTAAAGCTCCGTTCAACTCTGCTGCGGATGATATCTCCGAAACACTTGGTTGTTCT 438

QY 141 G1YTPAspI1eSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIysValPheAsp 160
DB 439 GGTGGATATATACAAACCGGATCTGTACGAAGCAATACACGTTCTTCAAGTTCTGGA 498
QY 161 I1eAspLeuGlnIYeuGlnLeuArgProTYrMetGluSerMetLeuProLeuProGlyI1e 180
DB 499 TATGATCTGCAACAGGCTGTGAAGCTAAGATGCTCTGTTAAGCACTGCCGTCCATC 558
QY 181 TYrAspProAspPheI1eAlaI1aAsnGlnGluIYrArgAlaAsnValI1e----- 198
DB 559 TACTACCCGATTTTATCGCACTAACCGACGAAACGTGTACAACTGATCAACCTG 618
QY 199 -----LysGlyThr-----LysGlnGluIYrValGlnI1eI1eLys 211
DB 619 GACGAAAGGTTAAGCTTACCTACCGGTGATGAGTCACTCACTGAGGTTACCGTCT 678
QY 212 AspI1eValAlaPheLysGluI1aThrIYsValAspLysValI1eValLeuTrpThrAla 231
DB 679 GATATCCAGAACTTCAAGAGAAACGCACTGACAAAGTTATCTGATCTGAGCTGCT 738
QY 232 AsnThrGluArgTYrSerAsnLeuValI1eGlyLeuAsnAspThrMetGluAsnLeuLeu 251
DB 739 AACACTGAAAGTTAGTGAAGTATCCCGGCTGTAAACGATTAAGTGAACAACTGCTG 798
QY 252 AlaAlaValAspArgAsnGluI1eGluI1eSerProSerThrLeuTYrAlaI1eLys 271
DB 799 CAATCTATCAAGACGACGACGAGAAATCGCTCCGTCACCATCTCCCTGCTGATCT 858
QY 272 ValMetGluAsnValProPheI1eAsnGlySerProGlnAsnThrPheValProGlyLeu 291
DB 859 ATCTCGAAGGCGTACCGGATCAACGCTCTCCGCAACACTTCTGATCCGAGCTG 918
QY 292 I1eAspLeuAlaI1eAlaArgAsnThrLeuI1eGlyIYrAspAspPheLysSerGlyGln 311
DB 919 GTACAGCTGCTGAAACAGCAAGTACCTTCACTCGCTGAGAGATCTGAATCTGCGCAG 978
QY 312 ThrIYsMetLysSerValLeuValAspPheLeuValGlyI1eLysProThrSer 331
DB 979 ACTTAACTGAAATCTGTAATCTGACACAGTTCCTGTTGACCGCTGTATCAACCGGTTCT 1038
QY 332 I1eValSerTYrAsnI1eLeuGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThr 351
DB 1039 ATCGCTCTTATACCACTGGGTAAACAGCGGTACCACTGCTGCTCGAAACAG 1098
QY 352 PheArgSerIYeuI1eSerIYsSerAsnValI1aAspAspMetValAsnSerAsnAla 371
DB 1099 TTCCGTTTAAAGAAATCTTAATCTCTGTATATCGACGATCATCGCTTAAACGAC 1158
QY 372 I1eLeuTYrGluPro-----GlyGluI1eProAspPheI1eValI1eLysTYrVal 389
DB 1159 ATCTGTTCAACGAAACCTGGAAGAGATCACTGTATCTTAAATCAATCAATG 1218
QY 390 ProTYrValGlyAspSerIYsArgAlaMetAspGlyTYrThrSerGluI1ePheMetGly 409
DB 1219 AAACCGTTGGTATCTTAAGTCTGTATGAGCAATCTACTCTGAAGTATGCTGGGC 1278
QY 410 GlyIYsSerThrI1eValI1eAsnThrCysGluAspSerLeuI1aAlaProI1e 429
DB 1279 GGTCAACAACCGATCTATATCAACAAGCTTTGTGAAGCTCTCTGCTGCTACCCGCTG 1338
QY 430 I1eLeuAspLeuValLeuLeuI1eGluLeuSerThrArgI1eGluPheLys----- 446
DB 1339 ATCATCGACCTGCTGTTATGACTAAATCTGTACCCGTTATCTTAAACAAAGTTGAC 1398
QY 447 -----AlaGluAsnGluI1eLysPheI1eSerPheI1eProValAlaThrI1eLeuSer 464
DB 1399 CCGGTTAAAGAGATGCTGCAAAATTCGAAACTTCAACCGGTTCTGACTCTTCTGCT 1458
QY 465 TYrLeuThrIYrAlaProLeuValProProGlyTYrProValI1aAsnAlaLeuSerLys 484
DB 1459 TACTGGCTGAAGGCTCCGCTGACTCGTCAAGCTTCCACCGGTTTAAACGCTGAACAA 1518

/ Sequence 12608, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949, 016
/ PRIORITY FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12608
/ LENGTH: 77626
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(77626)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12608

Alignment Scores:
Pred. No.: 3,07e-140 Length: 77626
Score: 1220.50 Matches: 311
Percent Similarity: 41.22% Conservative: 88
Best Local Similarity: 32.13% Mismatches: 105
Query Match: 46.37% Indels: 466
DB: Gaps: 12

US-10-718-952-2 (1-510) x US-09-949-016-12608 (1-77626)

QY 6 PhelyValaGlucyProbaenVallyrTyThrluThrluInserValTyraen 25
DB 32903 TTTCTGTGAGAGCCCGGACGTCGTCTACGCGCCGAGCGCATTCAGAGCGCATTCAGAG 32962
QY 26 TyrgluThrluThrluValuValuHsGlu----- 34
DB 32963 TACCGACGACGCGCTCAGCCGCGAGGCGCTTCTCAAGTGAACAGGCGGCTGAG 33022
QY 35 -----AsnArgAsnGlyThrTyrglnTrpIle 43
DB 33023 GAAGAGCGGTGCTGGAGAGCGGCGGCTCGGGAGAGAGGACCGGAGATT 33082
QY 44 -----VallyrProlyrSerVallyrTyrgluPhelyrThraen 56
DB 33083 CACTGCGCGGCTGTCTTCTGAGTGCACCCACGCTCCACGCGCTTCCCGGACCGCC 33142
QY 57 ILeHsValProlyrLeuGlyValMetLeuValGlyTrpGlyValAsnAsnGlySerThr 76
DB 33143 CGGCAAGTCCCGGCTCGGGGTCACTCTTGTGCGTGGCGGAGAACAGGCTCCACA 33202
QY 77 LeuThrglyValIleAlaAsnArgGlyIleSerTrpAlaThrlyrAsp----- 94
DB 33203 CTCACCGCGCGGTGCTGCACATCGATCGCTTGTCTGCGCCACGCGAGCGGCGCC 33262
QY 94 ----- 94
DB 33263 AAGTGGGGCGGAGCGGCGGTGGTGAAGAGCTTGTGGAGCTTCGCGGCGG 33322
QY 94 ----- 94
DB 33323 GGGCGGAGCGCTGCAAGGGGCGGGGCTCTGCGGACCAAGGGGCTGTAGGGGT 33382
QY 94 ----- 94
DB 33383 TCCGCGGGCGAGGGGCTGAGGGGCGAGGACCTCGCGGCTGTGAGGCTGTAGTG 33442

QY 94 ----- 94
DB 33443 GCGAGGACCTGACGAGGGCGAAAGCTGTGGGGCGGAGGGCTTCAGAGGGCGGGGCC 33502
QY 94 ----- 94
DB 33503 TGGCTCAGAGCGGACCCGTTGGAAGGTTTCGAGAGTACGGGGACTCTCGGAGTCT 33562
QY 94 ----- 94
DB 33563 TACAGGGCGGAGCTTAAGTGCAGAGGTCTCAGGGAGGGCTTGGGGCGAGCGCTT 33622
QY 95 -----LysIle----- 96
DB 33623 CGGTCTGCTGTGAGCTTGAACCTGAATAATGCCGAATGGGGCGGGCTTGAGAGT 33682
QY 96 ----- 96
DB 33683 CTAGGTGGGGCCGAGGCGGGTCCGGCCCTTTCACACCTCCACCCGTCACCTGCGCT 33742
QY 97 -----GlnGlnAlaAsnTyrrPheGlySerLeuThrluAlaSerAlaIleAr 112
DB 33743 GTCTGCTGACCCCGAGAGGCGCAACTACGCTGCTGACTCAGCGGCGACCGTGA 33802
QY 112 gValGly-----serPheGlnGlyGluGluIleTyrrAlaPropelyrSerLeuProme 131
DB 33803 CCGGCTGAGACCGCGAGGCGCGAGGCGTGTTCGTAACCTTCAGCGGTGCTGCCAT 33862
QY 131 tValAsnProAspAlaValPheGly----- 140
DB 33863 GGTGGGCGCCACAGACCTGTGTTCGA-TGTGTGGGCGAGCCCTGGCGGGGTGGGGG 33921
QY 140 ----- 140
DB 33922 GATGAGAGATGGGTCTGAGAGGCGCCAGATCCGCGGCGAGGCTGAGGCTGCC 33981
QY 141 -----GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaValPh 159
DB 33982 CCGAGGCTGGACATCTGCTGCTGACCTGACCGAGGCGGAGTCCGCGGAGGCTGCT 34041
QY 159 eAspIleAspLeuGlnGlyLeuAspProTyrrMetGluSerMetLeuProLeuProG 179
DB 34042 GAGCTGGGGCTGCGAGAGCAACTGTGGCGGACATGAGAGCGCTCGGCGCCGCTTC 34101
QY 179 yIleTyrrAspProAspPheIleAlaAlaGlnGluGluAlaAsnAsnValIlely 199
DB 34102 TGTTCATCTCCCGAATTCATCGCGGCGCAACAGAGCGCGCGGACCACTCATCCC 34161
QY 199 eGlyThrlyrGlnGluGlnVal----- 206
DB 34162 AGGCTGCGGTGCGAGAGGTCTGTCCATCCCGCATCCCTTGTCTGCGCACTGTT 34221
QY 206 ----- 206
DB 34222 CTTCGCACTTGCTTCTCTTTTGGCCCAACCCCGGCAATCCAGGCTTCATTCAGACC 34281
QY 206 ----- 206
DB 34282 CGGGAACTGATAGCCAGGCTCTGTGAGTCCGAGATATGTACCAACCACTTCCTA 34341
QY 207 -----GlnGlnIleIlely 211
DB 34342 CACCTCTTCATCTCCCAACACCTTCACCATCAACCAACAGCTGAGACATTCGCG 34401
QY 211 sAspIleIleValaPheLyserGluAlaThrlyrValaAspLyserValaValIleuTrpThrl 231
DB 34402 GGAATTCGAGACTTCGCTGAGCGGGGCTGGAACAAGTCATATGCTCTGAGAGCGC 34461
QY 231 sLeuThrglyTyrrSerAsnLeuValaIleLyserAsnAspThrMetGluAsnLeu 251
DB 34462 GAACACGAGGCGCTTCTGTGAGATTCAGGCTTCACAGCAGACAGCGAGAACTGCT 34521
QY 251 uAlaIleAlaValaAspArgAsnGlu-Ala----- 259

Db	34522	CGCAGCATTGAGTGGGCGAAGCCCGGGGTGAGCGCGGTGACGACGAGCGGCACGTACACA	34581
Oy	260	----- -----GluIleSerProSerT	265
Db	34582	CACAGCTCTGCGTGACTGCCTCCGCTTGCACAGCTCGGTCTGAGAGGTGCCTCTCA	34641
Oy	265	hrlLeuYrAlaIleAlaCysValMetGluAsnValProPheIleAsnGlySerProGlnA	285
Db	34642	CGCTCTTGCGCGGTGGCAGCATCTCTGAGGGCGTGTCTCTTCATAGGGTCTCCGACA	34701
Oy	285	snThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeuIleGlyIYA	305
Db	34702	ACACCTCGGTGGCCCGAGAGCTCTTGAGCTTCGCGTGGCAGCAGCCGGTCTTTTGGGCGGAG	34761
Oy	305	spApPheIleYsSerGlyGlnThrIlyMetIysSerValLeuValAspPheLeuValGlyA	325
Db	34762	ATGACTTCAGATCAGGCGCAGCAACMAAGTCAAGTCGCTTGCTTGACCTTCATATGGCT	34821
Oy	325	laGlyIleIlyS-----	328
Db	34822	CCGGCTCTCAAGGTGCGGTGGCCTTAGGAGCTGCCAGATGACGAGAAGGGGGCTCGGGCCGC	34881
Oy	328	-----	328
Db	34882	GAGCAGCTGTGGTGGGCGCGAGGCGCTGACAGCTTGGGGCTTCCTTGACCCACA	34941
Oy	329	--ProThrSerIleValSerTyrTrpAsnIleuGlyAsnAspGlyMetAsnLeuSerA	348
Db	34942	GACCATG-TCCATGTGATGATTAACAACACTGGGCAACAACGATGGGAGAACTATGCG	35000
Oy	348	laProGlnThrPheArgSerIleGlnIleSerIysSerAsnValValAspAspMetValA	368
Db	35001	CGCATATGCAGTTCCTCCGCTTAGAGAGGTGTCCAAGAGCAACGTGTGTGACGACATGTGTC	35060
Oy	368	snSerAsnAlaIleLeuYrGlnuProGlyIlyuHisProAspHisVal-----	383
Db	35061	AGAGCAACCCAGTGTCTATACGCCCGGCGCAAGCCTGACCACTG-CGTGCGTGGGGCG	35119
Oy	383	-----	383
Db	35120	CGGCGCGGGCGCGGGGTTGCCCGGCGAGGGGTGGCAGGAGCCCGGGCAAACTCTGCTG	35179
Oy	384	-----ValValIleIlySerTyrValProTyrValGlyAspSerIlyAsArgAlaMetAspG	401
Db	35180	CATCTCAGGTGTGATCAAGATATGTGCGTACGTGGTGTGACACCAAGCGCGCGCTGAGT	35239
Oy	401	IuYrTrpSerGluIlePheMetGlyGlyIysSerThrIleValLeuHisAsnThrCysG	421
Db	35240	AGTATACCTCGAGAGCTGATGCTGTGGCGGAACCAACACTGTGTGTGACACACCTGTG	35299
Oy	421	Iu-----	421
Db	35300	AGGTGCGGCGGCTTCAAGGGGCTGTATGGAGACAGGGGCTTGGCCACCGCCCATCT	35358
Oy	422	-----AspSerIleuLeuAlaIleProIleIleLeuAspLeuVal	434
Db	35360	CCTGACCCACCGCGCCGCGAGACTGCTGTGGCGCGCACCATCATATGCTGAGACTTAGG	35419
Oy	435	LeuLeuAlaGluLeuSerThrArgIleGluPheIlyAlaGluAsnGlnGlyIysPheHis	454
Db	35420	CTGTGTACCGAGCTGTGCCACCGGCTGAGGCTTTCGACCTACAGTGAAGCCCGGCGGAG	35479
Oy	455	SerPheHisProValAlaThrIleLeuSerTyrIleuThrIlyAsIaProLeuValProPro	474
Db	35480	ACCTTTCACCCCGTGTGTCTCTCTCTTCTTCAAGCGGCACCTAGTGTGCGGCC	35539
Oy	475	GlyThrProValValAsnAlaLeuSerIlyGlnArgAlaMetLeuGluAsnIle-----	492
Db	35540	GGCAGCCCGGTGTGTATATGCGCTTTTCCGACAGGCGAGCTGCATCGAGAACTCTCAGG	35599
Oy	492	-----	492

DB	Accession	Sequence	Length	Matches	Conservative	Mismatches	Indels	Gaps
OY	492	-----	1231	209	56	3	1	
DB	35660	TGCACCCAGCCTCCAGAGTCCGCCCGGGGCCAGATACCCCTCGCTAGTATGCA	35655					
OY	493	-----	1231	209	56	3	1	
DB	35660	CAGGCTTGGGGGCTGTGATGTCGAGGATACCCCTTCTCCCTGTGGGTCTCG	35719					
OY	493	-----	1231	209	56	3	1	
DB	35720	CTCTGCTAGCTGTGTGACCTCTCCAGGGGCTCTCGTGGGGCTCCGCCACAGAACCAT	35779					
OY	505	-----	1231	209	56	3	1	
DB	35780	GCTCTGGAACACAAA 35795						
RESULT 12								
US-09-397-787-34								
; Sequence 34, Application US/09397787								
; Patent No. 6468758								
; GENERAL INFORMATION:								
; APPLICANT: Benson, Darin R.								
; APPLICANT: Lodes, Michael J.								
; APPLICANT: Mitcham, Jennifer L.								
; APPLICANT: King, Gordon E.								
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN								
; FILE OF INVENTION: CANCER THERAPY AND DIAGNOSIS								
; TITLE REFERENCE: 210121.466C2								
; CURRENT APPLICATION NUMBER: US/09/397,787								
; CURRENT FILING DATE: 1999-09-16								
; NUMBER OF SEQ ID NOS: 334								
; SOFTWARE: PasteSeq for Windows Version 3.0								
; SEQ ID NO 34								
; LENGTH: 1231								
; TYPE: DNA								
; ORGANISM: Homo sapien								
US-09-397-787-34								
Alignment Scores:								
Pred. No.: 5.4e-128 Length: 1231								
Score: 1098.50 Matches: 209								
Percent Similarity: 81.29% Conservative: 56								
Best Local Similarity: 64.11% Mismatches: 58								
Query Match: 41.74% Indels: 3								
DB: 3 Gaps: 1								
US-10-718-952-2 (1-510) x US-09-397-787-34 (1-1231)								
OY	187	AlAlaLaSngIngluIarGAlaA-AsnAsnValIleYsgIYThrLYeGIngluIaVa	206					
DB	5	GGCGCCAAACCAAGACCGCGCGCGGACAAACCTCATCCAGGCTCGCGGACAGCT	64					
OY	206	IGIngluIleIleYAspIleYsAlaPheLYsguIaIaThrLYsValAspLYsAlVa	226					
DB	65	GGACCAATCCGACGGGACATCCGAGACTTCCGCTTACGCGCGGCTGAGCAAACTCAT	124					
OY	226	IVAlleuTriPThrAlaSnThrguIarqYrSerAsnLeuValIaIGLYeuaSnAspTh	246					
DB	125	AGTCTGTGACGGCGCAACCGAGCCCTTCTGTAGAGTGATTCAGGCTCAACGACAC	184					
OY	246	rMeGluAsnLeuLeuAlaIaValAspArgAsnGuIaIGluIeSerProSerThrie	266					
DB	185	AGCGGAAACCTGCTCGCACCATTAAGCTCGGT--CTGAGAGTGTCCCTCCACGCT	241					
OY	266	uTYrAlaIleAlaCySValMetGluAsnValProPheIleAsnGlySerProGluAsnTh	286					
DB	242	CTTGGCGGTGACCAAGCAATCTGAGAGGCTGTGCTCTCTCAAGAGGTCTCCGACGAC	301					
OY	286	rPheValProGluYleuIleAspLeuAlaIleAlaArgAsnThrLeuIleGlyLYsAsp	306					
DB	302	CTGTGCTCCCGGAGCTCTTAGCTCCGCTGCGACGACACCGGATTTTGTGGGCGAGATGA	361					
OY	306	pPheLYsSerGlyInThrLYsMeLYsSerValLeuValAspPheLeuValIGLYAlaGI	326					
DB	362	CTTCATGTACGGCGCAGACCAAGATCAAGTCCGCTGTGTGAGCTTCTCATAGGCTCCG	421					

```

QY 326 ylllelysProthSerileValSerTyraenhiSleuGlyAsnaAsnpglywMeAsn-1 346
Db 422 CCTCAAGACCATGTCATGATGATTAACAACCTGGCAACAGCGGAGAACCC 481
QY 346 euSerAlaProGlnThrPheArgSerLyseGluLeSerLyseSerAsnValAlaAspAspM 366
Db 482 TATCGCGCATTCGACGTTCCGCTTAAGAGAGGTGTCAAGAACACGTGTGAGACAA 541
QY 366 etValAsnSerAsnAlaileuTyrgluProGlyGluhiSProAspHiValVal11 386
Db 542 TGGTGCAGAGCAACCCAGTGTCTATACGCCCGGCAAGAGCTGACCACTGCGTGTCA 601
QY 386 lelyerTyraProTyrrValGlyAspSerLyseValArgAlaMetAspGlyTyrrThSerGlu 406
Db 602 TCAGATGATGCGCGTACGTGGTGCAGCAAGCGCGCTGATGATGATTAACCTCGAGAG 661
QY 406 laphemetGlyGlySerThrileValLeuhiSAsnThrCySGluAspSerLeuLeu 426
Db 662 TGAATGCTGGGCGGACCAACAACACTGTGTGCACAACAGTGTGAGACTCGCTGTGG 721
QY 426 lAlaProIleileuAspLeuValleuLeuAlaGluLeuSerThrArgIleGluPheL 446
Db 722 CCGCAACCATCATGCTGACCTGACCTGCTGACCGAGCTGTGCGAGCGGTGAGCTTCT 781
QY 446 yAlaGluLeuGluGlyLysePhehiSerPhehiSerProValAlaThrIleLeuSerTyrl 466
Db 782 GCACCTGACATGACCCCGAGCGCGACACTTCCACCCCGTGTCTGCTGCTGCTGCTCC 841
QY 466 euThrTyrrAlaProLeuValAlaProGlyTyrrProValAlaAsnAlaLeuSerLyseGlu 486
Db 842 TCTTCAAGCGCGCATTAAGTCCGCCCGGACAGCCGGTGTGCTTAATGCGCTTTTCCGAGC 901
QY 486 rGAlaMetLeuGluAsnIleMetArgAlaCyValAlaGlyLeuAlaProGluAsnAsnMet 506
Db 902 GGAGCTGCATCGAGACATCTCTCAAGGCGCTGCGTGGGCTCCCGCACAAGAACCATGCG 961
QY 506 lAlaGluTyrrLyse 510
Db 962 TCCTGGAAACAA 975

```

```

RESULT 13
US-09-902-540-1257/c
; Sequence 1257. Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1257
; LENGTH: 34316
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1257

```

```

Alignment Scores:
Pred. No.: 1.13e-55 Length: 34316
Score: 545.00 Matches: 145
Percent Similarity: 47.31% Conservative: 101
Best Local Similarity: 27.88% Mismatches: 204
Query Match: 20.71% Indels: 71
DB: 4 Gaps: 13
US-10-718-952-2 (1-510) x US-09-902-540-1257 (1-34316)

```

```

QY 17 GluThrGluIleGlnSerValTyraenTyrgluThrGluLeu-----Val 32
Db 20023 GAAATTTAAAGTATGAACCTTTAAGATATCAACATCCGCTTTCAGATTTTCCGCATG 19964
QY 33 H1sglu-----AsnaArgAsnGlyThr 39
Db 19963 CAGAACCGGTATCGTACCGACCGGCATACGGCGCATCGAATACGTCTAACGAAGAACAA 19904
QY 40 TyrgluThrIleVallyseProlyseSerVallyTyrgluPheThrAsnIleH1Val 59
Db 19903 TCACGATGGAACAAAGAGTGTGTCGCG-AGCGCGAAGCGGAAGCTCCGCTCGAAT 19845
QY 60 ProlyseGluGlyValMetLeuValGlyTyrglyAsnaAsnGlySerThrleuThrGly 79
Db 19844 CCGGCTGTGGGCGCTGTG-----TCAACAACGCTGATGGCG 19809
QY 80 GlyValIleAlaAsnaArgGluGlyIleSerTyrrAlaThrlyAspAspIleGlnAla 99
Db 19808 GGTGTGACCTGGCGCGCAAGGCG-----AGGGT 19779
QY 100 AsnTyrrPheGlySerleuThrGlnAlaSerAlaIleArgValGlySer--PheGlnGly 118
Db 19778 CACCCCATTTGGGTGCTGCTCAGCAGATGGGACCGCGCTGGGGAAGCGACCGAGCGG 19719
QY 119 GluGluIleTyrrAlaProPheLyseSerleuLeuPrometValAsnProAspAspIleVal 138
Db 19718 CCGACCGTG-----AGCTCAACGAGCTGGTGCCTTGTGGAGAGCTGAAGAGCTTGCC 19665
QY 139 PheGlyGlyTyrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleVal 158
Db 19664 TTCGGGCGCTGGGACATCATCCGTGAGACGCGATACGAAGTGGCGGTGCGTGGCGGTG 19605
QY 159 PheAspIleAspLeuGlnlyseGluLeuArgProTyrrMetGluSerMetLeuProleuPro 178
Db 19604 CTCAGGACCAACACCTCGAGAGAGGTGAAGCCGTTCTTCCAGAGCATCAAGCGGAAGCAG 19545
QY 179 GlyIleTyrrAspProAspPheIleAlaIleAsnGlnGluArgAlaAsnaAsnValIle 198
Db 19544 GGGGTGCAGACCTGAGTTCGTGGCGCGCATCGAGGCC-----AACCAATC 19497
QY 199 lYsglyThrLyse-----GlnGluGlnValGlnGlnIleIlelyseAspIleLyseAlaPhe 216
Db 19496 AAGCGCACCAACACGACCGCGAGAGCATCGAAGCGTGGCGGACGACATCCGCACTTC 19437
QY 217 lYsgluAlaThrLyseValAspLyseValValIleTyrrThrAlaAsnThrGluArgTyrr 236
Db 19436 AAGAAGAGCTCAACGCGACGCGCGCGGTGATGTGACGACGCTGAGACCTTC 19377
QY 237 SerAsnLeuValValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArg 256
Db 19376 CGTCCGCTGCCGAGTCTTCAAG--ACGTGGCGCGCTTGCAGAGAGCGCTGAGCGAG 19320
QY 257 AsnGluAlaGluIleSerProSerThrleuTyrrAlaIleAlaCyValMetGluAsnVal 276
Db 19319 AACAGCGCGGACATCAACCCACCGCGCTGTACCTAGCAGCGGCATCAAGAGGCGGTG 19260
QY 277 PropheileAsnGlySerProGln--AsnThrPheValProGlyLeuIleAspLeuAla 295
Db 19259 CCCTTGCAGAACGCGACCGCCCAAGCGCAGCGGTGACACCGCGCGCTGACGAGAGGCGC 19200
QY 296 lIleAlaArgAsnThrleuIleGlyLyseAspAspPheLyseSerGlyGlnThrlyMetLyse 315
Db 19199 AAGCAGAGCTCGGTGGCGCTCCGCGCGCGCATCTMAAGCGGCGACGATGAAGAAG 19140
QY 316 SerValleuValAspPheLeuValGlyAlaGlyIleTyrrProthSerileValSerTyrr 335
Db 19139 ACGGTATATGCGCCCGCGCTCAAGGCGCGACGCTGCGGTGATGAGCTGCTTCTCAC 19080
QY 336 AsnhiSleuGlyAsnaAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLyse 355
Db 19079 AACATCTTGGCAACCGCGACGCGAGGTGCTGACAGACCCGCGGCTTCAAGGCGCAAG 19020
QY 356 GluIleSerLyseSerAsnValValAspAspMetValAsnSerAsnAlaIleleuTyrglu 375

```


QY 460 AlaThrIleuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValVal 479
DB 1165 -----ATGCTCTTACTTCAAGACCCCAAGCCGCGGCTGCGGTGAG 1215
QY 480 AsnAlaLeuSerLysGlnArgAlaMetLeuGluAsnIleuMetArgAlaCysValGlyLeu 499
DB 1216 CACGACTCTTCACTCAGCTCACCAAGCTGAGACACGCTGCGGTGCGGTGAG 1275
QY 500 AlaProGluAsnMetIleuGlyTyr 509
DB 1276 GACCCCATCACCCTCGGCTCGACTAT 1305

RESULT 15

US-09-313-294A-4684
/ Sequence 4684, Application US/09313294A
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Laligudi, Raghunath V.
/ APPLICANT: Ito, Laura V.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313, 294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 4684
/ LENGTH: 294
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. 6476212 70034858H1
/ NAME/KEY: unsure
/ LOCATION: 185, 272, 275, 281, 283, 288
/ OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4684

Alignment Scores:

Pred. No.:	4.23e-10	Length:	294
Score:	155.00	Matches:	27
Percent Similarity:	80.43%	Conservative:	10
Best Local Similarity:	58.70%	Mismatches:	9
Query Match:	5.89%	Indels:	0
DB:	4	Gaps:	0

US-10-718-952-2 (1-510) x US-09-313-294A-4684 (1-294)

QY 1 MetPheIleGluAsnPhelValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 125 ATGTTATCGAAGCTTCGCGAGACCCCAAGCTGCGGTACGCGCCGAGGATC 184
QY 21 GlnSerValTyrAsnTyrGluThrThrGluValIleGluAsnArgAsnGlyThrTyr 40
DB 185 NTCTCGAGTACCGTACGACACGACGAGCTGTGTACAGGCCCAAGACGCGCTCC 244
QY 41 GlnTrrIleValLysPro 46
DB 245 CGCTGGTGTCTCGGCC 262

Search completed: June 7, 2005, 22:57:50
Job time : 310 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2005, 20:45:48 ; Search time 3675.5 Seconds
(without alignments)
860.161 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632

Sequence: 1 MREIKFKVCEPVMVXTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/US10718952/runat.06062005.173402.12940/app.query.fasta_1.1294
-DB=Published Applications NA -QMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR.SCORE=apct -THR.MAX=100
-THR.MIN=0 -ALLEN=15 -MODE=LOCAL -OUTFMT=pico -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USRR=US10718952.@CGN_1.1.879.@runat.06062005.173402.12940
-NCPU=6 -ICPU=3 -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US10A_NEW_PUB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2632	100.0	1533	14	US-10-025-003-9	Sequence 9, Appli
2	2632	100.0	1533	14	US-10-718-952-9	Sequence 9, Appli
3	2632	100.0	1760	14	US-10-025-003-1	Sequence 1, Appli
4	2632	100.0	1760	14	US-10-718-952-1	Sequence 1, Appli
5	2627	99.8	1533	14	US-10-025-003-5	Sequence 5, Appli
6	2627	99.8	1533	19	US-10-718-952-5	Sequence 5, Appli
7	2619	99.5	1589	14	US-10-424-599-70167	Sequence 70167, A
8	2600	98.8	1533	14	US-10-025-003-13	Sequence 13, Appli
9	2600	98.8	1533	14	US-10-025-003-15	Sequence 15, Appli
10	2600	98.8	1533	19	US-10-718-952-13	Sequence 13, Appli
11	2600	98.8	1533	19	US-10-718-952-15	Sequence 15, Appli
12	2593	98.5	1533	14	US-10-025-003-11	Sequence 11, Appli
13	2593	98.5	1533	19	US-10-718-952-11	Sequence 11, Appli
14	2524	95.9	1533	18	US-10-424-599-12022	Sequence 12022, A
15	2378	90.3	1533	9	US-09-938-842A-1438	Sequence 1438, Ap
16	2378	90.3	1533	11	US-09-938-842A-1438	Sequence 25858, A
17	2356	89.5	2156	18	US-10-425-114-25858	Sequence 47189, A
18	2356	89.5	2284	20	US-10-425-115-47189	Sequence 14714, A
19	2355	89.5	1924	19	US-10-767-701-14714	Sequence 2091, Ap
20	2354	89.4	1941	20	US-10-739-930-2091	Sequence 122151, A
21	2351	89.3	2444	20	US-10-425-115-122151	Sequence 70769, A
22	2347	89.2	2576	19	US-10-437-963-70769	Sequence 7081, Ap
23	2345	89.1	1898	18	US-10-425-114-7081	Sequence 10, Appli
24	2336	88.8	1931	9	US-09-921-232-10	Sequence 10, Appli
25	2336	88.8	1931	9	US-09-921-330-10	Sequence 5704, Ap
26	2336	88.8	1931	8	US-10-425-114-5704	Sequence 5612, Ap
27	2335	88.7	1879	18	US-10-425-114-5704	Sequence 1, Appli
28	2335	88.7	1914	18	US-10-425-114-5612	Sequence 26983, A
29	2335	88.7	1959	9	US-09-727-628-1	Sequence 47194, A
30	2335	88.7	1971	18	US-10-425-114-46983	Sequence 84893, Ap
31	2335	88.7	2233	20	US-10-425-115-47189	Sequence 47190, A
32	2335	88.7	2203	20	US-10-425-115-47189	Sequence 205, App
33	2334	88.7	1912	18	US-10-425-114-4683	Sequence 85285, A
34	2334	88.7	1997	20	US-10-425-115-47190	Sequence 14009, A
35	2283	86.7	1759	19	US-10-363-829-205	Sequence 5864, Ap
C 36	2205.5	83.8	2196	18	US-10-437-963-85285	Sequence 83553, A
37	2193	83.3	1853	18	US-10-425-114-16524	Sequence 47182, A
38	2193	83.3	1858	18	US-10-425-114-16524	Sequence 12021, A
39	2193	83.3	2050	18	US-10-425-114-1864	Sequence 12021, A
40	2193	83.3	2130	20	US-10-425-115-83553	Sequence 47180, A
41	2129	80.9	1673	20	US-10-425-115-47180	Sequence 47180, A
42	2104	79.9	1789	20	US-10-425-115-47182	Sequence 12021, A
43	2023	73.8	2582	18	US-10-424-599-12021	Sequence 15, Appli
44	1942.5	73.8	2582	18	US-10-424-599-12021	
45	1813	68.9	3546	9	US-09-921-232-15	

ALIGNMENTS

RESULT 1
US-10-025-003-9
; Sequence 9, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997

PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 9
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-025-003-9

Alignment Scores:
Pred. No.: 1.15e-263 Length: 1533
Score: 2632.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-718-952-2 (1-510) x US-10-025-003-9 (1-1533)

QY	1	MetPheIleGluAspPheValGluCysProAsnValIleTyrThrGluThrGluIle	20
DB	1	ATGTTCAATCGAAGATTAAAGTTGAGTCTCTTAATGTAAGTACACCGAAGCTAGATT	60
QY	21	GlnSerValTyrAsnTyrGluThrThrGluLeuValHieGluAsnArgAsnGlyThrTyr	40
DB	61	CAGTCGGTGTACACATACGAACACCGAATCTGTTACAGAGACAGAAATGGACCTAT	120
QY	41	GlnTrrIleValIlePyrProLysSerValIleTyrGluPheLysThrAsnIleHisValPro	60
DB	121	CAGTGAATGTCAAACCCAAATCTGTCAATACGAATTTAAACCAACATCATGTTCTCT	180
QY	61	LysIleuGluValMetLeuValGluTrrPglYglYAsnAsnGlySerThrLeuThrGluGly	80
DB	181	AAATTAGGGGTAAATGCTTGTGGGTGGGTGGAAACAAGGCTCAACCTCTCACGGTGT	240
QY	81	ValIleAlaAsnArgGluGlyIleSerTrrPAlaThrLysAspLysIleGlnGlnAlaAsn	100
DB	241	GTTATTGCTTAACCGAAGGGCATTTTCATGGGCTACAAAGACAAGTTTCAACAAGCAT	300
QY	101	TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu	120
DB	301	TACTTTGGCTCTCCCTCACCAAGCTCAGCTATCCGAGTGGGCTCTTCCAGGAGAGAA	360
QY	121	IleTrrAlaProPheLysSerLeuLeuProMetValAsnProAspIleValPheGly	140
DB	361	ATCTATAGCCCATTCAGAGCTGCTTCCAATGTTAACCTGACGACATGTGTGGG	420
QY	141	GluTrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp	160
DB	421	GGATGGGATATACGACAACTGAACCTGGCTGATGCGCATGGCCAGGCAAGGTGTGAC	480
QY	161	IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle	180
DB	481	ATCGATTTCAGAAAGCAGTTGAGGCTTCAATGGAATTCATGCTCCCGGGAATC	540
QY	181	TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly	200
DB	541	TATGACCCCGAATTTCAATGTGCGCAACCAAGAGCGTGCACAAAGCATCAAGGGGC	600
QY	201	ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr	220
DB	601	ACAAAGCAAGAGAAATTCACAAATCATCAAGACATCAAGCGCTTAAAGAGACAC	660
QY	221	LysValAspLysValValIleuTrrPAlaAsnThrGluArgTyrSerAsnLeuVal	240
DB	661	AAAGTGAACAAGGTGTTGACTGTGACTGCCAACACAGAGAGGTAAATTTGGTT	720
QY	241	ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu	260
DB	721	GTGGGCTTAAATGACCAATGGAATCTTGTGGCTGTGTGACGAATAGGGCTGAG	780
QY	261	IleSerProSerThrLeuTrrAlaIleAlaCysValMetGluAsnValProPheIleAsn	280

DB	761	ATTTCTCTCCATCACTTGTATGCCATTCCTGTGTATGAAATGTTCTTCATTAAT	840
QY	281	GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr	300
DB	841	GGAAAGCCCTCAGMAACATTTGTACCAAGGGCTGATGTGATCTTGCCATGCGAGGAAC	900
QY	301	LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp	320
DB	901	TTGATTTGTGAGATGATCTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTGTTGAT	960
QY	321	PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn	340
DB	961	TTCTTGTGGGGGCTGATATCAAGCAACATCTATAGTCACTTCAACCATCTGGGAAC	1020
QY	341	AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer	360
DB	1021	AATGATGGTATGAACCTTTCGGCTCCAAACCTTCCGTCCAGGAATCTCCAAAGAGC	1080
QY	361	AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro	380
DB	1081	AACGTTGTGATGATATGTCACACGCAATGCCATCTTATGAGCCTGGTGAATCA	1140
QY	381	AspHisValValIleLysTrrValProTyrValGlyAspSerLysArgAlaMetAsp	400
DB	1141	CACCATGTTGTGTTATTAAGATGTGCTTACGTAGGGGACAGCAAGAGCCATGAT	1200
QY	401	GluTrrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys	420
DB	1201	GAGTACACTTCAGAGATATTCATGGGTGAAAGACACCATTTGTTTGCAACAACATGC	1260
QY	421	GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer	440
DB	1261	GAGGATTCCTCTTTCGCTGCTCTTATCTTGACCTTGCTCTTCTTGCTGAGCTGAGC	1320
QY	441	ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla	460
DB	1321	ACTGAATCGAGTTTAAAGCTGAAATAGGGAATTCATCTCACCACCTGCTCT	1380
QY	461	ThrIleLeuSerTrrLeuThrLysAlaProLeuValProProGlyThrProValValAsn	480
DB	1381	ACCATCTTCAGTACTTCAACCAAGGCTCTCTGTGTTCCACGGGTACACCATGTGTGAT	1440
QY	481	AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla	500
DB	1441	GCATTGTCAAAGCAAGCTGCAATGCTGGAATAACATAAGAGGCTGTGTGATTGGCC	1500
QY	501	ProGluAsnAsnMetIleLeuGluTrrLys	510
DB	1501	CCAGAGAATAACATGATTCCTCGAGTACAAAG	1530

RESULT 2
US-10-718-952-9
Sequence 9, Application US/10718952
Publication No. US20040128713A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/718, 952
PRIORITY FILING DATE: 2003-11-21
PRIORITY APPLICATION NUMBER: 08/835, 751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 9
LENGTH: 1533

TYPE: DNA
ORGANISM: Glycine max
US-10-718-952-9

Alignment Scores:

Pred. No.:	1.15e-263	Length:	1533
Score:	2632.00	Matches:	510
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	19	Gaps:	0

US-10-718-952-2 (1-510) x US-10-718-952-9 (1-1533)

```

QY      1 MetPheIleGIuAnPheLySValIGuCyProAsnValLySfYrThrgIuThrgIuIle 20
DB      1 ATGTCATCGAGAAATTTTAAGTTGAGTGTCTTAAGTGAAGTCAACCGAGCTGAGATT 60
QY      21 GlnSerValTyraSnTyrgIuThrgIuLeuValHleGIuAnPrgAnSgIyThrTyx 40
DB      61 CAGTCCGTGTAACACTACGAAACACCAAGCTGTTCACGAAACAGAAATGGACACTTAT 120
QY      41 GlnTrpIleValLySProLySerValLySfYrGluPheLyThraSnIleHleSValPro 60
DB      121 CAGTGAATTGTCAAAACCAATCTGTCAATACGAATTTAAACCAACATCCATGTTTCT 180
QY      61 LysLeuGIyValMetLeuValIGlyTPGIyGIyAsnAnSgIySerThrLeuThrgIyGIy 80
DB      181 AAATTGGGGGTATATGCTGTGGGTGGGTAACCAACGGCTCAACCTCCACGGTGGT 240
QY      81 ValIleAlaAnPrgGIuGIyIleSerTPAlaThrLySAspLySfIleGlnIleAlaAn 100
DB      241 GTTATTGTCTAACCGAAGGGCAATTTCATGGCTTCAAAAGCAAGATTCACCAAGCCAAAT 300
QY      101 TyrPheGIySerLeuThrgIuAlaSerAlaIleArgValIGlySerPheGlnGIyGIu 120
DB      301 TACTTTGGGCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGCTCTTCCAGGAGAGGAA 360
QY      121 IleTyraIlePrgPheLySerLeuLeuProMetValAsnProAspAplIleValPheGIy 140
DB      361 ACTATAGCCCACTTCAAGAGCCGTCTTCCATGTGTTAACTTCAGCAATGTGTTGGG 420
QY      141 GlyTPAspIleSerMetLeuAlaAspAlaMetAlaArgAlaIleValPheAsp 160
DB      421 GGATGGGAATATCGAACAATGAACCTGGCTGATGCCATGGCCAGGCAAGGAGGTTTGGAC 480
QY      161 IleAspLeuGlnLySfIleuArgProTyMetGIySerMetLeuProLeuProGIyIle 180
DB      481 ATCGATTTGCAGAAAGCAAGTTGAGGCTTACATGGAATCCATGCTTCCATCCCGGAATC 540
QY      181 TyrAspProAspPheIleAlaAlaAnSgIuGIyArgAlaAnAnValIleLySgIy 200
DB      541 TATGACCCGGAATTTCAATGCTGCCAACAAGAGAGGTGCCAACAACGTCATCAAGGGG 600
QY      201 ThrLySgIuGIuGIuAlaGlnGlnIleIleLySAspIleLyValaPheLySgIuAlaThr 220
DB      601 ACAAAGCAAGAGAGAAATCAACAATCATCAAGACATCAAGGCGTTTAAAGAGGCAACC 660
QY      221 LysValaAspLySfValaValIleuTPThraIaAsnThrgIuTrgTySerAsnLeuVal 240
DB      661 AAAGTGAACAAGGTGTTGTACTGTGACTGCCAACAAGAGGAGTACGATATTTGGTT 720
QY      241 ValGIyLeuAnAspThrMetGIuAnLeuAlaAlaValaAspArgAnSgIuAlaGIu 260
DB      721 GTGGGCTTATATGACACCATGGAATCTCTTGGCTGTGAGCAAGAAATGAGGCTGAG 780
QY      261 IleSerProSerThrLeuTyraIleAlaIleAlaCysValMetGIuAnValProPheIleAsn 280
DB      781 ATTTCTCTTCCACCTGTGATGCAATGCTTGTGTATGGAATAATGTTCTTTCATTAAAT 840
QY      281 GlySerProGIuAnSnThrPheValProGIyLeuIleAspLeuAlaIleAlaArgAnThr 300
DB      841 GGAAGCCTCAGAACCTTTTGTACCAAGGCTGATGTGATCTTCCATCCGAGAGAACACT 900

```

```

QY      301 LeuIleGIyIleAspAspPheLySerGIyGlnThrLyMetLySerValLeuValaAsp 320
DB      901 TTGATTGGTGGAGAGACTTCAAGAGTGGTCAAGACCAAAATGAATCTGTGGTTGAT 960
QY      321 PheLeuValaGIyAlaGIyIleLySProThrSerIleValaSerTyraSnIleSgIyAAn 340
DB      961 TTCCTTGTGGGGCTGTGATTCAGCAACATCTATATGTCAGTTACCAACATCTGGGAAAC 1020
QY      341 AsnAspGIyMetAsnLeuSerAlaProGIuThrPheArgSerLySgIuIleSerLySer 360
DB      1021 AATGATGATGAATCTTCCGCTCCCAACACTTTCCTCCAGAGGAATCTTCAAGAGC 1080
QY      361 AsnValaValaAspAspMetValaSerAsnAlaIleLeuTyGIuProGIyGIuHleSPro 380
DB      1081 AACGTTGTGATGATATGGTCAACAGCAATGCAATCTCTATAGCCTGTGTAACATCA 1140
QY      381 AspHleValaValaIleLySfYrValProTyValaGIyAspSerLySfArgAlaMetAsp 400
DB      1141 GACCATGTTGTGTTATTTAAGTATGTGCTTACGTAGGGGACAGCAAGAGCCATGAT 1200
QY      401 GIuTyThrSerGIuIlePheMetGIyGIyLySfSerThrIleValaIleuHleSnThrCyS 420
DB      1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGACACCATTTGTTGCAACAACATGC 1260
QY      421 GIuAspSerLeuLeuAlaProIleIleLeuAspLeuValLeuLeuAlaGIuLeuSer 440
DB      1261 GAGGATTCCTCTTACGCTCTCTTATATCTTGAATGTGCTCTTCTTCTGAGCTCAGC 1320
QY      441 ThrArgIleGIuPheLySfAlaGIuAnSgIyLySfPheHleSerPheIleProValaIle 460
DB      1321 ACTAGATGAGTTTAACTGAAATGAGGAAATTCATCACTCATCCACAGTTGCT 1380
QY      461 ThrIleLeuSerTyLeuThrLySfAlaProLeuValProProGIyThrProValaIleAn 480
DB      1381 ACCATCTCAGCTACCTCAACAAGGCTCTCTGTTCCACCGGTACCAAGGTGGAAT 1440
QY      481 AlaLeuSerLyGIuArgAlaMetLeuGIuAnIleMetArgAlaCysValaGIyLeuAla 500
DB      1441 GCAATGTCTAAACAGGTCATGCTGGAACCATATATAGGCTGTGTGATTTGGCC 1500
QY      501 ProGIuAnAsnMetIleLeuGIuTyLyS 510
DB      1501 CAGAGATTAACATGATCTTCAGATACAG 1530

RESULT 3
US-10-025-003-1
; Sequence 1, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-1
Alignment Scores: 1.41e-263 Length: 1760

```

Score: 2632.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-718-952-2 (1-510) x US-10-025-003-1 (1-1760)

```
QY 1 MeCPheIIeGIuAsnPhelYsValGIuCyProAsnValIysTYrThrgIunhrgIuile 20
DB 54 ATGTCATCGAGAAATTTTAAGGTGAGTGTCTTAAGTGAAGTACCCAGACATGAGAT 113
QY 21 GInSeValTYrAsnTYrGIunhThrgIuEuValIeGIuAsnAgsnGIYhTYr 40
DB 114 CAGTCGCTGAACAATACGAAACCAAGACTTGTTCACGAGAACGAAATGCACTAT 173
QY 41 GInTPILeValIysProIysSeValIysTYrGIuPhelYsThrsnIleHISValPro 60
DB 174 CAGTGAATGTCAAAACCAATCTGTCAAAATGCAATTAACCAACATCAATGTTCT 233
QY 61 LysLeuGIValMeIleuValGIYTPGIYGIYAsnAsnGISeThrThrgIYGIY 80
DB 234 AAATTTGGGGTAATGTTGTGGGTGGGTGGAAACAAGCTCAACCTCAGCGGTGAT 293
QY 81 ValIleAlaAsnAgsnGIYIleSeTYrPalathrIysAspIysIleGIunIalaAsn 100
DB 294 GTTATGCTAACGAGAGGCAATTCATGGGCTACAAAGACAAAGATTCACCAAGCCAT 353
QY 101 TYrPhelYSeIleuThrgIunIalaSerAlaIleYrValGIYSeThrheGIunGIYGI 120
DB 354 TACTTGGTCTCCCTCAAGAGCTCAGCTATCCAGTTGGGTCTTCAAGGAGAGAA 413
QY 121 ILeTYrAlaProPhelYSeIleuProMetValAsnProAspIleValPheGIY 140
DB 414 ATCTATGCCCATTCACAGAGCTGCTTCCAATGTTAACCTGACACATGTGTTGGG 473
QY 141 GIYTYrAspIleSeThrsnMeIleuValAspAlaMeAlaThrgIalYsValPheAsp 160
DB 474 GGATGGGATATCGACAATGAACCTGGCTGATGGCATGGCCAGGCAAAAGGTGTGAC 533
QY 161 ILeAspLeuGIunIleuValrProTYrMetGIuSeMeIleuProIeuProGIYIle 180
DB 534 ATCGATTTTCAGAGAGAGTTGAGGCTTTCATGGAATTCATGTTCCACTCCCGGAATC 593
QY 181 TYrAspProAspPheIleAlaIleAsnGIunGIYAspIleYsAlaPheIysGIY 200
DB 594 TATGACCCGGATTTCTTGTCTGCCAACCAAGAGAGCTGCCAACACATCATCAAGGAC 653
QY 201 ThrlYsGIunGIunIleuValIleIleYsAspIleYsAlaPheIysGIunIathr 220
DB 654 ACAAAACCAAGAGCAATTCACAAATCATCAAAAGATCAACAGGCTTTAAGGAGCCACC 713
QY 221 LysValAspIysValIleValIleuThrPalathrIysGIuYrISeThrsnIleuVal 240
DB 714 AAAGTGAACAAGGTGTTTACTGTGACTGCCAACAGAGAGGTACGTAATTTGGT 773
QY 241 ValGIYLeuAsnAspThrMetGIuAsnIleuValAlaValAspAgsnGIunIagIu 260
DB 774 GTGGGCTTAAATGACCAATGGAATCTTGGCTGTGGACGAATGAGGCTGAG 833
QY 261 ILeSeProSeThrIleuTYrAlaIleAlaCYsValMeGIuAsnValProPheIleAsn 280
DB 834 ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGTTATGGAATAATCTTCTTCAATAT 893
QY 281 GIYSeProGIunAsnThrPheValProGIYLeuIleAspLeuAlaIleAlaAgsnThr 300
DB 894 GGAAGCCCTCAGAAACCTTTTTCACAGGAGGATGATTCCTCCATCCCGAGAACACT 953
QY 301 LeuIleGIYIYAspAspPheIysSeGIYGIunThrIysMeIleuSeValIleuValAsp 320
DB 954 TTGATTTGGTGAATACCTTCACAGAGTGTGACAAACCAAAATCTGTGTTGGTGTAT 1013
QY 321 PheIleuValGIYAlaGIYIleYsProThrSeIleValSeTYrAsnHISleuGIYAsn 340
```

```
DB 1014 TTCCCTTGGGGGCTGGATACAGCAACATCTAATAGTACAGTTACAAACATCTGGGAAAC
QY 341 AsnAspGIYMeAsnIleuSeValProGIunThrPheAgsSeTYsGIunIleSeTYsSer 360
DB 1074 AATGATGGTAAATCTTTCGGCTCCAAACCTTCCGTTCCAGGAATCTCCCAAGAC 1133
QY 361 AsnValIleAspAspMeValAsnSeThrsnAlaIleIleuTYrGIuProGIYGIunIasPro 380
DB 1134 AACGTTGTGATGATATGATGTCACAGCAATGCATCTCTATAGACCTGTGTGAATCCA 1193
QY 381 AspHISValValIleIleYsTYrValProTYrValGIYAspSeTYsAlaIleMeIAsp 400
DB 1194 GACCATGTGTTGTATTAAATGATGTGCTTAACTGAGGAGACAGCAAGAGCCATGAT 1253
QY 401 GIuTYrThrsnGIunIlePheMeGIYGIYIysSeThrIleValIleuHISAsnThrCYs 420
DB 1254 GAGTACACTTCAGAAATATTCATGSGGTGAAAGACACCATTTGTTGCAACAACATGC 1313
QY 421 GIuAspSeIleuValAlaProIleIleuAspLeuValIleuValAGIuIleuSer 440
DB 1314 GAGATTTCCCTCTTATGCTGCTCTATTAATCTTGAGCTTGCTCTTGTGAGCTGAGC 1373
QY 441 ThrATGIleGIuPhelYsAlaGIuAsnGIuGIYIysPheHISerPheHISProValAla 460
DB 1374 ACTAGAAATCGAGTTTAAAGCTGAAATGAGGAAATTCACCTCAATCCACCCAGTTCT 1433
QY 461 ThrlIleuSeTYrIleuThrIysAlaProIleuValProProGIYThrProValIAsn 480
DB 1434 ACCATCTCAGCTACTTCAACAGGCTCTCTGTTCCAGCGGTACACAGTGTGAT 1493
QY 481 AlaIleuSeTYsGIunAgsAlaMeIleuGIuAsnIleMeIAlaCYsValGIYleuAla 500
DB 1494 GCATTGTCAAAGCAGCGTGAATGCTGGAAAAACATTAAGAGGCTTGTGATTTGCC 1553
QY 501 ProGIuAsnAsnMeIleuGIunTYrIys 510
DB 1554 CCAGAGAAATTAACATGATTTCTGAGTACAAAG 1583
```

RESULT 4

US-10-718-952-1
; Sequence 1, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-1

Alignment Scores:

Score: 1,416-263 Length: 1760
Pred. No.: 2632.00 Matches: 510
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-10-718-952-2 (1-510) x US-10-718-952-1 (1-1760)

QY	1	MetPheIleGIyAaenPheIyValGIyCysProaenValIyVtYrThGluThrGIuThrGIuLe	20
Db	54	ATGTTTCATCGAAGAAATTTTAAGGTGAAGTCTTAATGTGAAGTACACCGAAGCTAGATT	113
QY	21	GIuSerValIYrAaenTyrGIuThrThrGIuLeuValIhISGIuAaenAraSangIYhTyr	40
Db	114	CAGTCCGTGTCAACCTACGAAACCAACCGAATTGTTCACGAGAACAGGAATGGCACTTAT	173
QY	41	GIuTrrPIIeValIyPProIySereValIySyrIYrGIuPheIyThrAenIleHISValPro	60
Db	174	CAGTGAATGTGCAAACCCAAATCTGTCAAAATCGAATTTAAACCAACATTCATGTCTCT	233
QY	61	LysIeuGIyValIeMetLeuValGIYrPogIyGIyAaenAaGISeTThrThrThrGIyGIy	80
Db	234	AAATTTAGGGGTAAATCTTGTGGGTGGGTGAAACACACGCTCAACCTCCACCGGTGT	293
QY	81	ValIleAaenAraGIuGIyIleSeTrrPalatrrIyAaenIyAaenIleGIuIuIaAaen	100
Db	294	GTTATTTGCTTAACCGAAGGGCATTTCAATGGCTTAACAAAGAACAAATTCACACGCAAT	353
QY	101	TyrPheGIySeriIeuThrGlnIaIaSerIaIaIeAraValGIySerPheGIuGIyGIuGIu	120
Db	354	TACTTTGGCTCCCTACCCAAAGCCTCAGCTATCCAGATGGGTCTTCCAGGGAAGGAA	413
QY	121	IleTrrAlaProPheIySereIeuIeuProMetValIaenProAaenAaPIIeValPheGIy	140
Db	414	ATCTAATGCCCATTCAGAGGCTGCTTCAATGGTTAAACCTGACAGCATTTGTGTGGG	473
QY	141	GIYrTrrAaPIIeSeriAaenMetAaenIuIaAaenIaAaenIaAaenIaAaenIaAaenIaAaen	160
Db	474	GGATGGGATATCAGACAAATGAACCTGGCTGAATGCCATGGCCAGGGAAGAGGTGTGGAC	533
QY	161	IleAaenIeuGIuIyGIuIeuAraProTrrMetGIuSeMetIeuProIeuProGIyIle	180
Db	534	ATCGATTTGCAGAAAGCACTTGAAGCCTTACATGAATCCATGCTTCCACTCCCCGGAATC	593
QY	181	TyrAaenProAaenPheIleAaIaAaenGIuGIuAraGIaAaenAaenValIleIyGIy	200
Db	594	TATACACCGGATTTTCATGTGTGCCAACCAAGAGGAGCGGCCAACAGTCATCAAGGAC	653
QY	201	ThrrIySengIuGIuIuValGIuGIuIleIleIyAaenIleIyAaenIleIyAaenIleIyAaen	220
Db	654	ACAAAGCAAGAGCAAGTTCACAAATCATCAAAAGCATCAAGGCGTTTAAAGAACCCAC	713
QY	221	IyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaen	240
Db	714	AAAGTGGCAAGAGGTGTGTACTGTGACTGTCACACAGAGGATCAAGTATTTGGTT	773
QY	241	ValGIyLeuAaenAaenPrrMetGIuAaenIeuIuIaIaValAaenAaenAaenIuIaGIu	260
Db	774	GTGGGCCCTTAATGACACCATGAGAAATCTTTGGCTGTGGACAGAAATAGGCTGAG	833
QY	261	IleSerProAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaen	280
Db	834	ATTTCCTCTTCAACCTGTATGCAATCTCTGTGTATGAAATAGTTCCTTCAATAT	893
QY	281	GIySerProGIuAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaen	300
Db	894	GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATTTGATCTTGCCATCGCGAGAACACT	953
QY	301	IeuIleGIyGIyAaenAaenPheIySereGIyGIuIyIyMetIySereValIeuValAaen	320
Db	954	TTGATGTGTGAAGTACTTCAAGAGTGTCAAGCAAAATGAATATGTGTGTGTAT	1013
QY	321	PheIeuValGIyAaenIyIleIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaen	340
Db	1014	TTCTTGTGGGGCGGTATCAAGCAACATTAATGTCAGTTAACACATCTGGAAGAAC	1073
QY	341	AaenAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaenIyAaen	360
Db	1074	AATATAGGTAAATCTTTGTGGCTTCCAAACCTTCCGTTCAGAGAAATCTCCAAAGAC	1133

OY	361	AsnValValAaPaaPaaMeValAaAsnSerAsnAlaIleuValTrGluProGluGlnHisPro	380
Db	1134	AACGTTGTATGATATGTGTACAAAGCAATGCCATCTCTATAGAGCCGTGGAAATGCCA	1193
OY	381	AspHisValValIleuValTrValProTrValGluAaSerLysValAmeAaP	400
Db	1194	GACCATGTGTGTATTAATTAATGTGGCTTAACTAAGGGGACACAAAGAGCCATGGAT	1253
OY	401	GluValTrThrSerGluIlePheMetGlyGlyLysSerThrIleValIleHisAsnThrCys	420
Db	1254	GAGTACACTTCAGAGATATTCATGAGTGGAAAGAGACCAATGTGTTGGCAACAACATGC	1313
OY	421	GluAaSerLeuLeuAlaIleProIleIleuAaPLeuValIleuLeuAlaGluLeuSer	440
Db	1314	GAGATTCCTCCCTTAGCTGCTCCATATATCTGTGAACTTGGTCTCTTTGGCTGACCTAGC	1373
OY	441	ThrValIleGluPheLysAlaGluAaAsnGluGlyLysPheHisSerPheHisProValAla	460
Db	1374	ACTGAATTCGAGTTTAAAGCTGGAAATAGAGGAAATATCCACTCATATCCACCAGTTGCT	1433
OY	461	ThrIleuSerTrpLeuThrLysValAProLeuValProProGluTrpProValValAsn	480
Db	1434	ACCATTCCTCAGCTACCTCCACCAAGGCTCCTGTGTTCCACCGGTACACCAGTGGTAAT	1493
OY	481	AlaLeuSerLysGluAaGluAmeLysGluAsnIleMetAaGluAaCysValGlyLeuAla	500
Db	1494	GCATTGTCAAAGCAGCGTGCATCTCTGAAAAACATATAGAGGCGCTGTGGTGGATTGGCC	1553
OY	501	ProGluAaAsnMetIleuGluTrpLys	510
Db	1554	CCAGAGATTAACATGATTTCTCGAATTACAG	1563

RESULT 5

US-10-025-003-5
; Sequence 5, Application US/10025003
; Publication No. US20030074685A1

```

GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Street, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID

```

Alignment Scores:	
Pred. No.:	3, 8e-263
Score:	2627.00
Percent Similarity:	99.808
Best Local Similarity:	99.808
Query Match:	99.814
DB:	14
Length:	153
Matches:	509
Conservative:	0
Mismatches:	1
Indels:	0
Gaps:	0

US-10-718-952-2 (1-510) x US-10-025-003-5 (1-1533

Qy I MetPheIleGIuAsnPhelysValGIuCysProAsnValLysTyrThrGIuThrcIuile 20
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 ATGTCATCCAGAACTTTAAGGTGAAGTGCTCTAATGCAAGTACACCGAGACTGAGATT 60

```

QY 21 GlnSerValTyrAsnTyrGlnThrGlnLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGGGTACACATGACAAACCAACGAACTTGTTCAAGAACAGGAATGGCACTTAT 120
QY 41 GlnTTPLeuValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
DB 121 CAGTGGATGTGTAACCAACCAAACTGTCAATACGAATTTAAAACCAACATTCATGTTCT 180
QY 61 LysLeuGluValMetLeuValGlyTTPGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 181 AAATTTAGGGGTAAATGCTTTGGGTTGGGTTGGGTTGGAAACACAGCTCAACCTTCACCGGTGT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTTPAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATTGCTTAACCGAAGGCAATTTTCATGGCTCAACAAAGACAAAGATTCAACAGCAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTGGCTCCCTCCACCAAGCTTCAGCTATCCGAGTGGGTCTTCCAGGAGAGAGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCCCTTCAAGACCTGCTTCCAAATGTTTAACTTCAGACATTTGTGTTGGG 420
QY 141 GlyTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGCCAGGCAAAAGGTGTGGAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTTTCAGAAAGCAGTTGAGGCTTTCATGGAATTCATGCTTCACCTCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 601 ACAAAAGCAGAGCAAGATTCAACAAATCATCAAAAGATCATCAAGCGTTTAAAGAGCCACC 660
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 221 LysValAspLysValValIleuTTPThrAlaAsnThrGluArgTyrSerLeuLeuVal 240
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu 260
DB 721 GTGGGCTTAATGACCACTGAGAAATCTCTTGCTGCTGTGACCAAGAAATGAGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTGCTCTTCCACCTTGATGCAATGCTTGTTGTTATGGAATAATGTTCTTTCATTAAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAAGCCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGCCATCCCGAGAACACT 900
QY 301 LeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTGGTGGAGATGACTTCAAGATGGTCAAGCAAAATGAAATCTGTGTTGTTGAT 960
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 961 TTCCTTGTGGGGCTGTATCAAGCCCAACATCTATAGTCACTTACCAACCATCTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1021 AATGAGGTATGATCTTTTCGGCTCCCAAAACCTTCCGTTCCAAAGAAATCTCCCAAGAC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1081 AACGTTGTTGATGATATGCTCAACAGCAATGCCATCTCTATGAGCTGGTGAACATCCA 1140
QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400

```

```

DB 1141 GACCATGTTGTTGTTATTAAGATGTGCTTACAGAGGAGACAGCAATAGAGCCATGAT 1200
QY 401 GlnTyrThrSerGlnIlePheMetGlyGlyLysSerThrIleValIleuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAAATTTTCACTGGGTGGAAGAGACACATTTGTTTGCAACACATCC 1260
QY 421 GluAspSerLeuLeuAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTCTTATGCTGCTCCATTAATCTTGAGCTTGTCCTTCTTGCTGAGCTCAGC 1320
QY 441 ThrArgIleGluPheLysAlaGluAsnGlyGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAAGATCGAGTTTAAAGCTGAAATAGAGGAAATTCACCTCATTTCCACCACTTGCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB 1381 ACCATCTCAGCTTACTTCAACAGGCTCTCTGTTCCACCGGATKACACAGTGTGAAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1441 GCATTGTCAAAGCAGCGCAATGCTGGAAAAACATTAATGAGGGCTTGTTGATGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGlyTyrLys 510
DB 1501 CCAGAGAAATACATGATTTCTGAGTACAAAG 1530

RESULT 6
US-10-718-952-5
; Sequence 5, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-5

Alignment Scores:
Pred. No.: 3.8e-263 Length: 1533
Score: 2627.00 Matches: 509
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.81% Indels: 0
DB: 19 Gaps: 0

US-10-718-952-2 (1-510) x US-10-718-952-5 (1-1533)
QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 1 ATGTTCAATCGAAGAAATTTTAAGGTGAGTCTTAATGTGAATACCAACGACCTAGATT 60
QY 21 GlnSerValTyrAsnTyrGlnThrGlnLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGGGTACACATGACAAACCAACGAACTTGTTCAAGAACAGGAATGGCACTTAT 120
QY 41 GlnTTPLeuValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60

```

Db 121 CAGTGGATTGTCAAAACCAATCTGTCAAAATCAAAATTTAAACCAATCATGTTCTT 180
 QY 61 LysLeuGlyValMetLeuValGlyTTPGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
 Db 181 AAAATTAGGGGTAAATGCTTGGGGTGGGGTGAACCAACGGCTCAACCTCCACGGGTGGT 240
 QY 81 ValIleAlaAsnArgGlyGlyIleSerTPAlaThrLysAspLysIleGlnIleAlaAsn 100
 Db 241 GTTATTGCTAACCGAGAGGGCATTTTCATGGGCTCAAAAGACAAAGATTCAACAGCCAT 300
 QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGln 120
 Db 301 TACTTGGCTCCCTCCACCAAGCCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGGAA 360
 QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
 Db 361 ACTTAAGCCCATTCAGAGCCCTGCTTCCAAATGGTTAACCTTCAGCAATGTGTTGGG 420
 QY 141 GlyTPAlaAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
 Db 421 GGATGGGATATTCAGCAATCAATGAACTGGCTGATGCCATGGCCAGGCGAAGGTGTTGAC 480
 QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyIle 180
 Db 481 ATCGATTTCAGAGAGCAGTTGAGGCTTTCATGGAATCCATGCTTCCATCCCGGGAATC 540
 QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGlyLysArgAlaAsnAsnValIleLysGly 200
 Db 541 TATGACCCCGAATTTCTGCTGCTCCCAACAGAGAGACGTCGCAACACGTCATCAAGGCG 600
 QY 201 ThrLysGlnGlnGlnAlaGlnGlnIleIleLysAspIleLysValPheLysGlyAlaTyr 220
 Db 601 ACAAGCAAG 660
 QY 221 LysValAspLysValValLeuTyrThrAlaAsnThrGlyLysArgIleSerAsnLeuVal 240
 Db 661 AAAGTGCAGACAGGTGGTGTGATGCTGAGCTGCCAACACAGAGAGAGAGATTTGGTT 720
 QY 241 ValGlyLeuAsnAspThrMetGlyAsnLeuLeuAlaValAspArgAsnGlyAlaGlu 260
 Db 721 GTGGGGCTTATGACCACTGAGAAATCTTGGCTGCTGAGAGAGAAATGAGGCTGAG 780
 QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGlyAsnValProPheIleAsn 280
 Db 781 ATTTCTCTTCCACCTTGTATGCAATGCTTGTGTATGGAATATGCTTCTTCAATTAAT 840
 QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
 Db 841 GGAAGCCCTCAGAACACTTTTGTACAGGGGCTGATTTGATCTTGGCCATCCGAGAGAACCT 900
 QY 301 LeuIleGlyLysAspAspPheLysSerGlyInThrLysMetLysSerValLeuValAsp 320
 Db 901 TTGATTGGTGGAGATACCTTCAAGAGTGGTCAGAACCAAAATCTGTGTTGGTGAAT 960
 QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnIleLeuGlyAsn 340
 Db 961 TTCCTTGTGGGGGCTGGTATCAAGCCCAACATCTATATGATGATCAACACATCGGGAAAC 1020
 QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyIleSerLysSer 360
 Db 1021 AATTGATGATGATATTTCCGCTCCACAACTTTCCTTCCAGAGAAATCTCCAGAGAC 1080
 QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlyLysProGlyLysPro 380
 Db 1081 AACGTTGTGATGATATGCTCAACAGCAATGCAATCTCTATAGAGCTGATGAACTCA 1140
 QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
 Db 1141 GACCATGTGTTGTTATTAAGATATGCTTACGTAAGGAGACAGCAATAGAGCCATGAT 1200
 QY 401 GlyTyrThrSerGlyLysPheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
 Db 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAGACCAATCTTTTGCACAAACATGC 1260

QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuValLeuSer 440
 Db 1261 GAGGATTCCTCTTACTCTCTCTCTATATCTTGAACTGATGCTTCTTGTGAGCTCAGC 1320
 QY 441 ThrArgIleGlyPheLysAlaGlnAsnGlyLysPheHisSerPheHisProValAla 460
 Db 1321 ACTAGAAATGAGTTTAACTGAAAATGAGGGAAAATTCACATCTTCCACCCAGTTGGT 1380
 QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyTyrProValValAsn 480
 Db 1381 ACCATCTCAGCTACCTCACCAGAGCTCTCTGTTCCACCGGGTACACAGTGGTGAAT 1440
 QY 481 AlaLeuSerLysGlnArgAlaMetLeuGlnAsnIleMetArgAlaCysValGlyLeuAla 500
 Db 1441 GCAATGTCAAGCAGAGGTGCAATGCTGGAACCAATATAGAGGGCTTGTGTGATTTGGCC 1500
 QY 501 ProGluAsnAspMetIleLeuGlyTyrLys 510
 Db 1501 CCAAGAAATACATGATTTCTCGAGTACAG 1530
 RESULT 7
 US-10-424-599-70167
 ; Sequence 70167, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 70167
 ; LENGTH: 1989
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1
 US-10-424-599-70167
 Alignment Scores:
 Pred. No.: 3,81e-262 Length: 1989
 Score: 2619.00 Matches: 510
 Percent Similarity: 99.80% Conservative: 0
 Best Local Similarity: 99.80% Mismatches: 0
 Query Match: 99.51% Indels: 1
 DB: 18 Gaps: 0
 US-10-718-952-2 (1-510) x US-10-424-599-70167 (1-1989)
 QY 1 MetPheIleGlyAsnPheLysValGlyCysProAsnValLysTyrThrGlnThrGluIle 20
 Db 217 ATGTCATCGAAGAAATTTTAAGGTATGATGCTCTTAATGTGAAGATACCAAGACTGAGATT 276
 QY 21 GlnSerValTyrAsnTyrGlyThrThrGlnLeuValHisGluAsnArgAsnGlyThrTyr 40
 Db 277 CAGTCCGTACACTACCAACCAACCAACTTGTTCACGAGAAACGAAATGGCACTTAT 336
 QY 41 GlnTPAlaValLysProLysSerValLysTyrGlyLysPheLysThrAsnIleHisValPro 60
 Db 337 CAGTGGATTGTCAAAACCAATCTGTCAAAATGCAATTTAAACCAACATTCATGTTCTT 396
 QY 61 LysLeuGlyValMetLeuValGlyTTPGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
 Db 397 AAATTAGGGGTATATCTTGTGGGTGGGGTGAACCAACGCTCTACCTCAGCCGGTGG 456
 QY 80 ValIleAlaAsnArgGlyGlyIleSerTPAlaThrLysAspLysIleGlnIleAlaAsn 100
 Db 457 TGTATTGCTAACAGAGAGGCAATTTTCATGGGCTACAAAGACAAAGATTCAACAGCCAA 516

QY 100 nTyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 517 TTAACCTTGAGCTCCCTCACCAGCCACAGCTACAGTATCCAGTTGGGCTCTTCACAGGAGAGA 576
QY 120 uIleTyrAlaProPheLeuSerLeuLeuProMetValAsnProAspAspIleValPheG 140
Db 577 AATCTATGCCCAATTCMAAGAGCCTGCTTCCAAATGGTTAAACCTTGACAGCAATGTGTGG 636
QY 140 yGlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAs 160
Db 637 GGGATGGGATATCAGAACATGAACCTGGCTGATGTCATGCGCAAGGCAAGAGTGTTTAA 696
QY 160 pIleAspLeuGlnLysGlnLeuArgProTyrMetGlySerMetLeuProLeuProGlyT 180
Db 697 CATCGATTTGGAGAGACAGTGTGAGGCTTACATGGAATCCATGCTTCCATCCCGGGAAT 756
QY 180 eTyrAspProAspPheIleAlaAsnGlnGlyGluArgAlaAsnAsnValIleLysG 200
Db 757 CTATGACCCCGGATTTTCATGTGCTGCCAACAGAGAGGCTGCCCAACAGCTCATCMAAGG 816
QY 200 yThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysValPheLysGlnLanTh 220
Db 817 CACAAAGCAAGAGCAAGTTCACAAATCAACAAAGACATCAAGGCTTTAAGAGAGCCAC 876
QY 220 rLysValAspLysValValIleLeuTyrThrAlaAsnThrGluArgTyrSerAsnLeuVa 240
Db 877 CAAAGTGGACAGAGGTGGTGTGACTGTGAGCTGCCAACACAGAGAGGTATACGATTTGGT 936
QY 240 lValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaAlaValAspAspAsnGluAgl 260
Db 937 TGTGGGCTCTTAATGACACATGAGATCTCTTGGCTGTGGACAGAAATGAGGCTCA 996
QY 260 uIleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAs 280
Db 997 GATTTCTCTTCACCTGTATGCCATTTGCTGTGTATGAGAAATGTTCTTCTTTAA 1056
QY 280 nGlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnTh 300
Db 1057 TGGAAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATTCCTTGCCATGGCGAGAAC 1116
QY 300 rLeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAs 320
Db 1117 TTTGATTTGGTGGAGATGACTTCMAAGATGCTCAGACCAAAATGMAATCTGTGTGTTAA 1176
QY 320 pPheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHiIleGluGlyAs 340
Db 1177 TTTCTTTGGGGGCTGGATTCAGCCACATCTATATGATGATTCACACCATGTGGGAAA 1236
QY 340 nAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSe 360
Db 1237 CATATGATGGATTAATCTTTCCGCTCCACAACTTCCGTTCCAGAAATCTCCAAAGAG 1296
QY 360 rAsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyGluHiIsp 380
Db 1297 CAACGTTGTGTGATGATGATGCTCAACAGCAATGCTCATGACCTGCTGTAACATCTC 1356
QY 380 oAspHiIleValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAs 400
Db 1357 AGACCAATGTTGTATTAATGATGATGCTTACGTAAGGAGACAGAGAGAGCCATGGA 1416
QY 400 pGluTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHiIleAsnThrCy 420
Db 1417 TAGATACACTTCAGAGATATTCATGAGTGGAGAAAGACACATGTTTTCACACACATG 1476
QY 420 sGluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLysSe 440
Db 1477 CCGAGGATTCCTCTGTAGCTGCTCTATATCTTGGACCTGTGCTCTGCTAGCTCAG 1536
QY 440 rThrArgIleGluPheLysAlaGluLysGlnGlyLysPheHiIleProValAla 460
Db 1537 CACTAAGATCGATTAAAGCTGAAAAATGAGGAAAAATTCCTCACTCATCCCAAGTTGC 1596

QY 460 aThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAs 480
Db 1597 TACCATCTCAGCTACCTCACCAGGCTCTCTGTTCACCGGGGTACACCACTGGTGA 1656
QY 480 nAlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAl 500
Db 1657 TGCATGTCMAAGCAGCGGTGATCTGGAACATTAATGAGGCTTGTGTGGATTGGC 1716
QY 500 aProGluAsnAsnMetIleLeuGluTyrLys 510
Db 1717 CCCAGAGATATAATCATGTTCTGAGTACAG 1747

RESULT 8
US-10-025-003-13
; Sequence 13, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-13

Alignment Scores:
Pred. No.: 2,476-260 Length: 1533
Score: 2600.00 Matches: 503
Percent Similarity: 99.41% Conservative: 4
Best Local Similarity: 98.63% Mismatches: 3
Query Match: 98.78% Indels: 0
Gaps: 0
DB: 14

US-10-718-952-2 (1-510) x US-10-025-003-13 (1-1533)

QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
Db 1 ATGTTCAATCGAAGAAATTTTAAGATAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGAT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 61 CAGTCGGGTACAACTACGAAACCAACGAACTTGTTCAGAGAAACAGAAATGGACATAT 120
QY 41 GlnTyrIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
Db 121 CAGTGGATGTGCAAAACCAATCCGTCACATCAATTAACCAACCAACCATGTGCA 180
QY 61 LysLeuGlyValMetLeuValGlyTyrPglGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db 181 AATATGGGGGTATGCTGTGGGTGGGAGAAACACGCGCTTACCTCAACCGGTGTGT 240
QY 81 ValIleAlaAsnArgGluGlyLysSerTyrAlaThrLysAspLysIleGlnGlnAlaAsn 100
Db 241 GTTATATGTCMAAGAGAGGAGGATTTCAAGGGCTTACAAAGAGAAATTCACAGCCAA 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGln 120
Db 301 TACTTTGGGCTCCCTCAACCCAGGCTCAGGATATTCAGATTTGATCTCTCAAGGAGAGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140


```

Db      361 ATCTAAGCCCATCTCAAGAGCTGCTTCCAAATGTTATCTCAACATCTGCTTGGG 420
Qy      141 GYTPAsp11eSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIysValPheAsp 160
Db      421 GGATGGGATATACGAACAACGACCTGGCTGATGCCATGGCCAGGGCAAGGTTGTTGAC 480
Qy      161 ILeAspLeuGlnIlySerGlnLeuArgProTyrMetGlnSerMetLeuProLeuProGlyIle 180
Db      481 ATCGATTTCGAGAGAGAGTTGAGGCGCTTACATGGAATCCATGGTTCCACCTCCCGGAATC 540
Qy      181 TYrAspProAspPheIleAlaAlaAsnGlnGlyIlyArgAlaAsnAsnValIleIysGly 200
Db      541 TAGACACCGGATTTCTATGCTGCTCCCAACAGAGAGCGTCCCAACAGCGATTAAAGGC 600
Qy      201 ThrIlyGlnGlnGlnValGlnGlnIleIleIysAspIleIysAlaPheIlySerGlnAlaThr 220
Db      601 ACAAGCCAAAGACAGCAATTCAGCAAAATCATCAAGACATCAAGGCGTTTAAAGAACCCACC 660
Qy      221 LysValAspLysValValIleuTyrThrAlaAsnThrGlnLysArgTyrSerAsnLeuVal 240
Db      661 AAAGTGAACAAGGTGTGTGCTGCTGAGACTGCAACACAGAGAGGTATAGCAATTTGGTT 720
Qy      241 ValGlyLeuAsnAspThrMetGlnAsnLeuAlaAlaValAspArgAsnGlnAlaGly 260
Db      721 CTAAGGCTTATATACACCATGAGCAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
Qy      261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGlnAsnValProPheIleAsn 280
Db      781 ATTTCTCTTCCACCTTGTATGCAATTCCTGCTGTGTATGGAATAATGTTCTTTCAATTAAT 840
Qy      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      841 GGAAGCCCTCAGAACACTTTGTACAGAGGCGTATGATCTTCCATCGGAGAGAACACT 900
Qy      301 LeuIleGlyIysAspAspPheIysSerGlyGlnThrIlyMetIysSerValLeuValAsp 320
Db      901 TTGATTGGTGGAGATATGACTTCAAGAGTGTGCACACCAAAATGAAATCTGTGTTGGTGGAT 960
Qy      321 PheLeuValGlyAlaGlyIleIysProThrSerIleValSerTyrAsnIleGlnIlyAsn 340
Db      961 TTTCTTGTGGGGGCTGATCAAGCCCAACATCTATATGTTATGTAACAACATCGGGAAAC 1020
Qy      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIlySerGlyIleSerIysSer 360
Db      1021 AATGAGGTATGATATCTCGGCTCCCAAAACCTTCGCCCAAGAGAAATCTCCAAAGAGC 1080
Qy      361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyIlyIleAspPro 380
Db      1081 AACGTTGTTGACATATGCTCAACAGCAATGCCATCTCTATAGCGCTGTGAAACATCCC 1140
Qy      381 AspHisValValIleIysTyrValProTyrValGlyAspSerIlyAspArgAlaMetAsp 400
Db      1141 GACCAATGTTGTATTAAGTATGCTTACGTAGGGGATAGCAAGAGCCATGGAT 1200
Qy      401 GlyTyrThrSerGlnIlePheMetGlyIlyIysSerThrIleValIleuHisAsnThrCys 420
Db      1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAAACCAATGTTTTCACCAACACATGT 1260
Qy      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGlnLeuSer 440
Db      1261 GAGGATTCCTTTTACCTGCTCTATTAATCTTGAATCTGCTCTTGTGTAAGCTGAGC 1320
Qy      441 ThrArgIleGlnPheIysAlaGlnIlyAsnGlnGlyIlyPheHisSerPheHisProValAla 460
Db      1321 ACTAGAAATCCAGTTTAAAGCTGAATAATGAGGAAATTCACATTCACCCAGTTGGCT 1380
Qy      461 ThrIleLeuSerTyrIleThrIlyAsnAlaProLeuValProProGlyThrProValIleAsn 480
Db      1381 ACCATTCTAGCTATATGACCAAGGCTCTCTGTGTCCACCGGTACACACAGGTGGGAAT 1440
Qy      481 AlaLeuSerIlyGlnArgAlaMetLeuGlnAsnIleMetArgAlaCysValGlyLeuAla 500

```

```

Db      1441 GCATTGTCAAGACGCGTGCATATGCTGGAAGAAATATATGAGGCTTGTGTGATTGGCC 1500
Qy      501 ProGlnAsnAsnMetIleLeuGlnTyrIys 510
Db      1501 CCAGAGAAATACATGATTCCTCGAGTACAAG 1530

RESULT 9
US-10-025-003-15
; Sequence 15, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Street, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIORITY FILING DATE: 2002-05-07
; PRIORITY FILING DATE: 08/835,751
; PRIORITY FILING DATE: APRIL 8, 1997
; PRIORITY APPLICATION NUMBER: PCT/US98/06822
; PRIORITY FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-15

Alignment Scores:
Pred. No.: 2,47e-260 Length: 1533
Score: 2600.00 Matches: 503
Best Local Similarity: 99.41% Conservative: 4
Query Match: 98.63% Mismatches: 3
DB: 98.78% Indels: 0
Gaps: 0

US-10-718-952-2 (1-510) x US-10-025-003-15 (1-1533)
Qy      1 MetPheIleGlnAsnPheIysValGlyCysProAsnValIysTyrThrGlnThrGlnIle 20
Db      1 ATGTTATCATGAAGATTTTAAAGTATGAGACTCTTAATGTAAGTACACCGAGCTGAGATT 60
Qy      21 GlnSerValTyrAsnTyrGlnThrThrGlnLeuValHisGlnAsnArgAsnGlyThrTyr 40
Db      61 CAGTCCGTTTACAACTACGAAACCAACGAACTTTGTTACGAGAAACGAAATGGCACTTAT 120
Qy      41 GlnTyrIleValIlyAspProIysSerValIysTyrGlnPheIysThrAsnIleHisValPro 60
Db      121 CAGTGGATGTCMAAACCCAAATCCGTCACATCACTTAATTAACCAACCAACCATGTTCCA 180
Qy      61 LysLeuGlyValMetLeuValGlyTyrGlyIysAsnAsnGlySerThrLeuThrGlyIly 80
Db      181 AAATTGGGGGTATGCTGTGGGTGGGTGGAACCAAGCGCTTACCTCAACCGGTGTGT 240
Qy      81 ValIleAlaAsnArgGlnGlyIleSerTyrAlaThrIlyAspIlyIleGlnGlnAlaAsn 100
Db      241 GTTATTGCTTACAGAGAGGCAATTCATAGCGCTTACAAAGACAAAGATTTCAACAGCAAT 300
Qy      101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlnIly 120
Db      301 TACTTGGCTCCCTCAACCAAGCTCAGCTATTTCAGATGTGATCTTCCAGGAGAGGAA 360
Qy      121 IlyTyrAlaProPheIysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db      361 ATCTATGCCCCATTAAGAGTGTGCTTCCAAATGTTAAATCCAGAGCAATTTGTGTTGGG 420
Qy      141 GYTPAsp11eSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIysValPheAsp 160
Db      421 GGATGGGATATACGAACAACGACCTGGCTGATGCCATGGCCAGGGCAAGGTTGTTGAC 480

```

161 ILeaPLeuGlnLysGlnLeuArpProTyrMetGluSerMetLeuProLeuProGlyIle 180
481 ATCGATTGCAAGACAGTGAAGCCTTACATGGAATCCATGTTCCATCCCGGAATC 540
181 TyrAspProAspPheIleAlaIaAngInGluIuArgAlaAsnAsnValIleLysGly 200
541 TACGACCCCGATTTCATGCTCTGCAACCAAGAGAGCGGCCAACACCTGATTGAAGGCG 600
201 ThrLysGlnGluValGlnGlnIleIleLysAspIleLysValPheLysGlnAlaThr 220
601 ACAAAGCAAGACAGATTGACAAATCATCAAGACATCAAGCCGTTTAAGGAAGCGAC 660
221 LysValAspLysValValIleLeuThrAlaAsnThrGluArgLysSerAsnLeuVal 240
661 AAAGGGAAGAAGGTGGTCTCTGTCGACTGCAACACAGAGAGGATGCAATTTGGTT 720
241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaIleValAspArgAsnGluAlaGlu 260
721 GTAGGCTTAATGACACCATGAGAAATCTTGGCTGCTGAGACAGAAATGAGGCTGAG 780
261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
781 ATTTCCTTCCACCTTGATGCAATGCTGTGATGGAATAATGTTCTTTCATTAAAT 840
281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTGATCTTGCCATCCGAGGAACAAT 900
301 LeuIleGlyValAspAspPheLysSerGlyInThrLysMetLysSerValLeuValAsp 320
901 TTGATTGGTGGAGATGACTTCAAGAGTGGTCAAGACCAAAATCTGTGTGGTGGAT 960
321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisIleuGlyAsn 340
961 TTTCCTGTGGGCTGTGATCAAGCCAAATCATAGTATGATACCAACATCGGGAAAC 1020
341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnLysSerLysSer 360
1021 AATGATGATGATGATCTCGGCTCCAAACCTTCGCTCCAGGAATACTCCCAAGAAC 1080
361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyLysPro 380
1081 AACGTTGTGAGATATGCTCAACAGCAATGCCATCTCTATGAGCCTGTGAAACATCCC 1140
381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
1141 GACCATGTGTGATTAAGATGCTTACGTAAGGGAATGCAAGAGCATGGAT 1200
401 GlyTyrThrSerGlnLysPheMetGlyLysSerThrIleValLeuHisIleAsnThrCys 420
1201 GAGTACACTTCAGAGATATTCAATGGGTGGAAGAACACCATGTTTGGCAACACATGT 1260
421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
1261 GAGGATTTCCCTTTAGCTCTCTATTAATCTTGACATTTGGTCTTCTCTGAGCTGAGC 1320
441 ThrArgIleGluPheLysAlaGluAsnGluLysPheHisSerPheHisProValAla 460
1321 ACTAGATCCAGTTTAAGCTGAAATGAGGAAATTCACATTCATTCACCCAGTTGGT 1380
461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
1381 ACCATTCTCAGCATTTGACCAAGGCTCTCTGTTCCACCGGATACACAGGTGGAT 1440
481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
1441 GGATTGCTCAAGAGCGCTGCAATGCTGAAACATATAGAGGCTGTGTGGATTGGCC 1500
501 ProGluAsnAsnMetIleLeuGluTyrLys 510
1501 CCAGAAATTAACATGATTTCTCGAGTACAG 1530

RESULT 10
US-10-718-952-13
; Sequence 13, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-13
Alignment Scores:
Pred. No.: 2,476-260 Length: 1533
Score: 2600.00 Matches: 503
Percent Similarity: 99.41% Conservative: 4
Best Local Similarity: 98.63% Mismatches: 3
Query Match: 98.78% Indels: 0
Gaps: 0
DB: 19
US-10-718-952-2 (1-510) x US-10-718-952-13 (1-1533)
QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 1 ATGTTTCATGCAAGAAATTTTAAGGTAGAGAGTCTCTAATGTAAGATACCGAGACTGAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCCGCTGATCACTACGAAACCAACCACTGTTCAAGAAACAGGAATGCACTAT 120
QY 41 GlnTPrIleValIleProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
DB 121 CAGTGGATTGTCAAAACCAATCCGTCAACTACCAATTTAAACCAACCAATGTTCA 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyLysAsnAngLysSerThrLeuThrGlyIle 80
DB 181 AAATTGGGGGTGATGCTTGGGGTGGGAGTGAACCAACGCGCTTACCTCACCGGTGGT 240
QY 81 ValIleAlaAsnArgGluGlyLysSerThrAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATTGCTCAACAGAGAGGCAATTCATAGGCTTACAAAGAACAAATTCACACAGCCAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluIle 120
DB 301 TACTTTGGCTCCCTACCCCAAGCCCTCAGCTATTCAGATTGGATCTTCCAGGAGAGGA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCAATCAAGAGTCTGCTTCCAAATGTTAAATCCTGACGACATTTGTGTTGG 420
QY 141 GlyTTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGGATATCAAGAAATGAACTGGCTGATGCAATGGCCAGGCAAGGTTGTTGAC 480
QY 161 ILeaPLeuGlnLysGlnLeuArpProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTGCAAGACAGTGAAGCCTTACATGGAATCCATGTTCCATCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaIaAngInGluIuArgAlaAsnAsnValIleLysGly 200

```

Db      541 TAGACCCGGATTCTTCTGCTCCAAACAGAGAGCGCTCCCAACACGATGATTAAGGC 600
Qy      201 ThrGlnGlnGlnValGlnGlnIleIleValAspIleValAlaPheValaThr 220
Db      601 ACAAAGCAAGACAGATTGACCAATCATCAAGACATCAAGCGTTTAAAGAACACCC 660
Qy      221 LysValAspLysValValIleLeuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db      661 AAAGTGAACAGGTGGTGTCTCTGTGACTGCAACACAGAGAGGTATAGCAATTTGGTT 720
Qy      241 ValGlyLeuAsnAspThrMetGluAsnLeuValAlaValAspArgAsnGluAlaGlu 260
Db      721 GTAGGCTTATACACCACTGAGGAATCTTGGCTGCTGTGACAGAAATGAGCTGAG 780
Qy      261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db      781 ATTTCTCTTCCACCTTGTATGCAATTCGCTGTGATGGAATAATGTTCTTTCATTAAT 840
Qy      281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      841 GGAAGCCCTCAGAAACCTTTGTACAGAGGCGTGAATGATCTGGCAATCCGAGAGAACCT 900
Qy      301 LeuIleGlyValAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db      901 TTGATTGGTGGAGTACCTTCAAGAGTGGTCAAGCCAAATGAAATCTGTGTTGGTTGAT 960
Qy      321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnIleGluIleAsn 340
Db      961 TTTCTTGTGGGGCTGTGTCAAGCCCAACATCTATAGTTACCAACCATCTGGGAAAC 1020
Qy      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleGluIleSerLysSer 360
Db      1021 AATGATGATGATGAATCTCGCTCCAAACCTTCCTCCCAAGAAATCTCCCAAGACC 1080
Qy      361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluIleAsp 380
Db      1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGTAACATCCC 1140
Qy      381 AspPheValValIleLysTyrValProTyrValGlyAspSerIleValAlaMetAsp 400
Db      1141 GACCATGTTGTGTATTAAGTATGTCCTTACGTAGGGGATAGCAAGAGCATGGAT 1200
Qy      401 GlnTyrThrSerGluIlePheMetGlyLysSerThrIleValLeuIleAsnThrCys 420
Db      1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCAATGTTTTCACAAACATGT 1260
Qy      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db      1261 GAGGATTCCTTTTACCTCTCTCTATTAATCTTGAGCTTGCTCTTCTGAGCTGAGC 1320
Qy      441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheIleSerPheIleAspValAla 460
Db      1321 ACTAGATTCAGATTAAAGCTGAAATAGAGGAAATTCACATCTTCCACCATGGTGT 1380
Qy      461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db      1381 ACCATTCTCAAGTATCTGACCAAGGCTCTCTGTTCCACCGGTACACACAGGGGTGAAT 1440
Qy      481 AlaLeuSerLysGluArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db      1441 GCATTGTCAAGACAGCGTGCATGCTGGAAACATATAGGGCTGTGTGGATTGGCC 1500
Qy      501 ProGluAsnAsnMetIleLeuGlnTyrLys 510
Db      1501 CCAGAGAAATACATGATCTCGAGTACAG 1530

```

```

; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strelitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 15
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-15

Alignment Scores:
Pred. No.: 2,47e-260 Length: 1533
Score: 2600.00 Matches: 503
Percent Similarity: 99.41% Conservative: 4
Best Local Similarity: 98.63% Mismatches: 3
Query Match: 98.78% Indels: 0
Gaps: 0
DB:
US-10-718-952-2 (1-510) x US-10-718-952-15 (1-1533)

Qy      1 MetPheIleGluAsnPheLysValGlyCysProAsnValIleTyrThrGluThrGluIle 20
Db      1 ATGTCATCAGAAATTTTAAGGTAGAGCTCTTAATGTGAATACCAAGACCTGAGATT 60
Qy      21 GlnSerValTyrAsnTyrGluThrThrGluLeuValIleGluAsnArgAsnGlyThrTyr 40
Db      61 CAGTCGCTGTACAACTACGAACCAACCAACTGTTTTCAGAAACAGATGGCACTAT 120
Qy      41 GlnTrpIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
Db      121 CAGTGATGTCGAACCAACCAATCCGTCAACTCAATTTAAACCAACCAACCATGTTCCA 180
Qy      61 LysLeuGlyValMetLeuValGlyTyrGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db      181 AATTTGGGGGTATGCTGTGGGTGGGTGGAACAACGGCTTACCTCAACCGGTGT 240
Qy      81 ValIleAlaAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
Db      241 GTTATTCCTACAGAGAGGGCTTTCAATGCGCTTACAAAGACAAAGATTCAACAGCCAT 300
Qy      101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db      301 TACTTTGGCTCCCTCACCAAGCTCAGCTATTCAGATTGATGATCTTCCAGGAGAGGAA 360
Qy      121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db      361 ATCTATGCCCATTTAAAGAGTGTCTTCCAAAGGTAAATCCCGAAGACATATGTGTGGG 420
Qy      141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db      421 GGATGGATATACGAACATGAACTGGCTGATGCGCAGGCAAGAGTGTGTTGAC 480
Qy      161 IleAspLeuGlnLysGluLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
Db      481 ATCGATTTCAGAAACAGATTGAGGCTTACATGAATCATGATGTTCCACTCCCGGAATC 540
Qy      181 TyrAspProAspPheIleAlaIleAsnGlnGluValArgAlaAsnValIleLysGly 200
Db      541 TAGACCCGGATTCTTCTGCTCCAAACAGAGAGCGCTCCCAACACGATGATTAAGGC 600
Qy      201 ThrLysGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
Db      601 ACAAAGCAAGACAGATTGACCAATCATCAAGACATCAAGCGTTTAAAGAACACCC 660

```

```

RESULT 11
US-10-718-952-15
; Sequence 15, Application US/10718952
; Publication No. US2004012873A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William

```

```
QY 221 LysValAspLysValValValLeuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGGACAAAGTGGTGTCTCTGTCGACTGCAACAGAGGATATGCAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAglu 260
DB 721 GTAGAGCTTAATGACCACTGAGGAATCTCTTGCTGCTGTGAGACGAATATGAGCTGAG 780
QY 261 IleSerProSerThrLeuTrpAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCTCTCCATCCATCTGTATGCCATGCTGTGATGAAATGTTCTCTTCAATTAAT 840
QY 281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATTTGATCTTGCCATCGCGAGAACACT 900
QY 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTGGTGGAGATGACTTCAAGAGTGGTCAGACCAAAATGAAATCTGTGTGGTTGAT 960
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 961 TTTCTGTGGGGGCTGGATCAAGCCACATCTAATGTTACACCACTTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyLysSerLysSer 360
DB 1021 AATGATGATGATATCTCGGCTCCAGAACCTTCCGCTCCAGAGAAATCTCCAGAAC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1081 AACGTTGTTGAGATGATGTCACACGAATGCCATCTCTATGAGCTGTGTGAAACATCCC 1140
QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCATGTTGTTGTTATGATGATGCTTACGTAAGGGAATGACAGAGGACCATGAT 1200
QY 401 GluTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAACACCATTTGTTGCAACACATAT 1260
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAgluLeuSer 440
DB 1261 GAGGATTCCTCTTAGCTCTCTATATCTTGGAACCTTGCTCTTCTGCTGAGCTGAC 1320
QY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCACATTCATCCACCAAGTTGCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB 1381 ACCATTCTCAGCATATGACCAAGGCTCTCTGTTCCACCGGTACACCAAGTGGAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1441 GGAATTCCTCAAGAGAGCTGCAATGCTGGAATAATGAGGCTTGTGTGATTTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGlyTyrLys 510
DB 1501 CCAGAAATPAACATGATTTCTCGAGTACAAAG 1530

RESULT 12
US-10-025-003-11
; Sequence 11, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
```

```
; CURRENT APPLICATION NUMBER: US/10/025.003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-11

Alignment Scores:
Pred. No.: 1,32e-259 Length: 1533
Score: 2593.00 Matches: 502
Percent Similarity: 99.22% Conservative: 4
Best Local Similarity: 98.43% Mismatches: 4
Query Match: 98.52% Indels: 0
DB: 14 Gaps: 0

US-10-718-952-2 (1-510) x US-10-025-003-11 (1-1533)

QY 1 MetPheIleGluAsnPheLysValGluCysProAsnValLysTyrThrGluThrGluIle 20
DB 1 ATGTTTCATCGAAGATTTTAAAGTAGAGAGTCTCTAATGTGAAGTACACCGAGCTGAGATT 60
QY 21 GlnSerValLysAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGGTGTACACTAGAACCAACCACTTGTTCAGAGAACAGAAATGGACCTTAT 120
QY 41 GlnTrpIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
DB 121 CAGTGGATTGTCAAAACCAATCCGTCAACTACCAATTTAAACCAACACCATGTTCCA 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyLysAsnGlySerThrLeuThrGlyGly 80
DB 181 AAATTGGGGGTGATGCTTGTGGGTGGGAGGAAACAACGGCTTACCTCCTCAGCTGGT 240
QY 81 ValIleAlaAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATTGCTAACAGAGAGACATTTCAATGGCTTACAAAGACAAATTCACACCACT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTTGGCTCCCTACCCAGACCTCAGCTAATTCAGATTGATCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValIleGly 140
DB 361 ATCTATGCCCTCAATCAAGAGTCTCTTCCAAATGTTAAATCTTGACGACATGTTGTGGG 420
QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGAATGGATATCAGAACATGAACCTGCTGATGTCAGGCGCAAGGAGGATTTGATC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
DB 481 ATCGATTTGCAAGAACAGATGAGGCTTACATGATTCATGATTCATCTCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 541 TACGACCCGGAATTTCAATGCTGCTCCAAACAGAGGAGCGTCCACAAAGTGTAAAGGCG 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
DB 601 ACAAAGCAAGACCAATTCAGAAATCATCAAAAGCATCAAGGCTTTTAAAGAACACAC 660
QY 221 LysValAspLysValValValLeuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGGACAAAGTGGTGTCTCTGTCGACTGCAACAGAGGATATGCAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAglu 260
```

```

Db 721 GTAGGCTTATGACACCATGGAATCTTGGCTGCTGGACGAATGAGGCTGAG 780
Qy 261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluValProPheIleAsn 280
Db 781 ATTTCTCTTCCACCTTGTATGCAATGCTGCTGTGTATGAAATGTTCTTCATTAA 840
Qy 281 GlySerProGlnAsnThrPheValProGlyLeuIleAsnLeuAlaIleAlaAsnThr 300
Db 841 GGAGGCCCTCAGAACCTTTGTACAGGGCTGATGATCTTGCATCGCAGAGAACT 900
Qy 301 LeuIleGlyValAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 901 TTGATTTGGTGAATGACTTCAGAGAGTCACAGCCAAATGAAATGCTGCTGTTGAT 960
Qy 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerThrAsnIleGluGlyAsn 340
Db 961 TTTCTGTGGGGCTGGTATCAAGCCAACTATAGTTAGTACAAACCATCTGGGAAAC 1020
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db 1021 AATGATGGTATGATCTCTGGCTCCACAAACCTTCGCTCAAGAAATCTCCAGAGC 1080
Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluIlePro 380
Db 1081 AACGTTGTTGACATATGCTCAACAGCAAGCCATCTCTATGAGCTGATGACATCCC 1140
Qy 381 AspHisValValAlaIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db 1141 GACCATGTTGTTGTTAAATGATGTCCTTACGTAGGGATAGCAAGAGCCATGAT 1200
Qy 401 GluTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
Db 1201 GAGTACACTTCAGAGATTCATGAGGGTGGAAGAACCACTGTGTTGCAACACATGT 1260
Qy 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1261 GAGGATTCCTTTTACTGCTGCTCTATTAATCTTGACATGCTCTCTCTGCTGAGCTGAGC 1320
Qy 441 ThrArgIleGluPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
Db 1321 ACTAGATTCAGCTTAAAGCTGAAATGAGGGAAATTCACATTCATCCACCATGTCCT 1380
Qy 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db 1381 ACCATTCAGCTATCTGACCAAGGCTCTGCTGCTCCACCGGCTACACCAAGTGAAT 1440
Qy 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1441 GCATTTCTCAAGACGCTGCAATGCTGGAACCAATTAATGAGGCTTGTGATGGCC 1500
Qy 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
Db 1501 CCAGAAATPACATGATTTCTCGAGTACAG 1530

```

```

; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-11
Alignment Scores:
Pred. No.: 1,32e-259 Length: 1533
Score: 2593.00 Matches: 502
Percent Similarity: 99.22% Conservative: 4
Best Local Similarity: 98.43% Mismatches: 4
Query Match: 98.52% Indels: 0
DB: 19 Gaps: 0
US-10-718-952-2 (1-510) x US-10-718-952-11 (1-1533)
Qy 1 MetPheIleGluAsnThrPheValGlyCysProAsnValLysThrGluThrGluIle 20
Db 1 ATTTCTATGAGAAATTTAAAGTAGAGACTCTTAATGTAGATACCGAGACTGAGATT 60
Qy 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 61 CAGTCCGTGACACATACCAACCACTGTTCAAGAGAACAGATGSCACTAT 120
Qy 41 GlnTrpIleValLysProLysSerValLysTyrGluPheLysThrAsnIleHisValPro 60
Db 121 CAGTGAATGTCAAAACCAATCCGTCAACTCAATTTAAACCAACCCATGTTCCA 180
Qy 61 LysLeuGlyValMetLeuValGlyTyrGlyValAsnAsnGlySerThrLeuThrGly 80
Db 181 AAATGGGGGTGATGCTTGTTGGTTGGGTGGGAAACAGAGCTTACCTCACCGGTGT 240
Qy 81 ValIleAlaAsnArgGluGlyIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
Db 241 GTTATTCATACAGAGAGACATTTTCATGGGCTGACAAAGACAAAGATTCAACAGCAAT 300
Qy 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 301 TACTTTGGCTCCCTCAACCAAGCTTCAGCTATTCGAGTTGATCTTCCAGGGAGAGAA 360
Qy 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 361 ATCTATGCCCATTCAGAGCTGCTTCCAAATGTTAACTTCAGCAACATGTGTTGG 420
Qy 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 421 GATGGGATATGACAACTGAACCTGCTGATGCCATGGCCAGGCAAGGCTTTGAC 480
Qy 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
Db 481 ATGATTTTCAGAGAGAGGTTGAGGCTTACATGAAATTCATGATGTTCCATCCCGGAATC 540
Qy 181 TyrAspProAspPheIleAlaAsnGlnGluValAsnValIleLysGly 200
Db 541 TACGACCCGGAATTCATGCTGCTCCCAACAGAGAGCGTGCCAAACCTGATTAAGAGGC 600
Qy 201 ThrLysGlnGluValGlnGlnIleLysAspIleLysAlaPheLysGluAlaThr 220
Db 601 ACMAAGCAAGACCAATGACCAATCACTCAAGACATCAAGCGCTTAAAGAGCCACC 660
Qy 221 LysValAspLysValValIleuThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db 661 AAAGTGAACAAGTGGTTCTCGTGAAGCTGCACACAGAGAGATGATGCAATTTGTT 720
Qy 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaValAspArgAsnGluAlaGly 260
Db 721 GTAGGCTTAAATGACCAATGAGATCTTGGCTGCTGTGAGCAATATGAGCTGAG 780
Qy 261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 781 ATTTCTCTTCCACCTTGTATGCAATGCTGCTGTGTATGAAATGTTCTTCATTAA 840

```

```
QY 281 GYSerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
    |||
Db 841 GGAAGCCCTCAGACACTTTTGACAGGCGTATTGATCTTGCCATCCCGAGGAACAAT 900
QY 301 LeuIleGlyValAspAspPheIysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
    |||
Db 901 TTGATTGGTGGAGATGACTTCAAGAGTGGTCAGAACCAAAATGTAATCTGTGGTGGAT 960
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
    |||
Db 961 TTTCCTGTGGGGCTGGATTCAGACCAACATCTAAGTTAGTACACCACTCGGGAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyLysSerLysSer 360
    |||
Db 1021 AATGATGATGATGATCTTCGGCTCCAAACCTTCGGCTCCAGAGAAATCTCCAAAGAC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyGlnHisPro 380
    |||
Db 1081 AACGTTGTGACGATATGGTCAACACCAATGCCATCTCTATGAGCCTGGTGAACATGCC 1140
QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
    |||
Db 1141 GACCATGTTGTTGTTATTAAGTATGTCCTTACGTAAGGAGTACCAAGAGCCATGAT 1200
QY 401 GlyTyrThrSerGlyIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCys 420
    |||
Db 1201 GAGTACACTTCAGACGATATTCATGGGTGGAAGAACACCATTTGTTGCAACAACATAT 1260
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuValLeuSer 440
    |||
Db 1261 GAGGATTCCTCTTACCTCTCTATTAATCTTGGAATGGTCTTCTCTGAGCGAGAC 1320
QY 441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
    |||
Db 1321 ACTAGATACAGTTTAAGTGAAGTGAAGGAAATTCACATCATTCACCCAGTTGCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyTyrProValValAsn 480
    |||
Db 1381 ACCATTCTCAGCATCTGACCAAGGCTCTCTGTTCCACCGGATCACACAGTGGAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
    |||
Db 1441 GCATTGTCAAGCAGCGTGCATGCGAAGAACATATGAGGGCTGTGTGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGlyTyrLys 510
    |||
Db 1501 CCAGAGATTAACATGATTCCTCGAGTCAAG 1530

RESULT 14
US-10-424-599-12022
; Sequence 12022, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12022
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110863C.1
US-10-424-599-12022
```

Alignment Scores:

```
Pred. No.: 3,066-252 Length: 2018
Score: 2524.00 Matches: 483
Percent Similarity: 98.24% Conservative: 18
Best Local Similarity: 94.71% Mismatches: 9
Query Match: 95.90% Indels: 0
DB: 18 Gaps: 0

US-10-718-952-2 (1-510) x US-10-424-599-12022 (1-2018)
QY 1 MetPheIleGluAsnPheLysValGlyCysProAsnValLysTyrThrGlnThrGlyIle 20
    |||
Db 98 ATGTTCAATCGACAGATTTCAAGTTAGAGTCTTAACGTGAATGACACAGAGCTGAGAT 157
QY 21 GlnSerValTyrAsnTyrGlnThrThrGlnLeuValHisGluAsnArgAsnGlyThrTyr 40
    |||
Db 158 CAGTCGGGTACCAACTATGAAGAACACTGAATCTTGTTCAAGAGAACAAATATATGCACTTAT 217
QY 41 GlnTyrIleValLysProLysSerValLysTyrGlyLysPheLysThrAsnIleHisValPro 60
    |||
Db 218 CAGTGGGTGTCCACCCCAAACTGTCAAAATATGAATTTAAACCAACCACTATGCTCT 277
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyValAsnAsnGlySerThrLeuThrGlyGly 80
    |||
Db 278 AACCTAGGGGTAAATCTGTGGGCTGGGGGGAACAAATGCTCAACCTCTCATGCTGTGT 337
QY 81 ValIleAlaAsnArgGluGlyIleSerTyrAlaThrLysAspLysIleGlnGlnAlaAsn 100
    |||
Db 338 GTTATTTGCTAACCGAGGGATTTTCATGGGCACAAAGAGACAAATTCACCAACCAAT 397
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGln 120
    |||
Db 398 TACTTGGGTCTCTCAACCGACATCAGCTATCAAGAGGGGTCTTTCACAGGGGAAGA 457
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
    |||
Db 458 ATATATGCTTCCATTCAAGAGCTGCTTCCATGAGAACCTCGATGATCTGTGTGGG 517
QY 141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
    |||
Db 518 GGATGGGATATCAGTAACTTGAACCTGGCTGATGCAATGGCCAGGGCCAAAGGTGTTGAT 577
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetLeuProLeuProGlyIle 180
    |||
Db 578 ATCGACCTGCACAAACAGTTGAGCGGTAACATGAAATCATGATCTCCCGGAATC 637
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluArgAlaAsnAsnValIleLysGly 200
    |||
Db 638 TATGACCCGGATTTTCATTTGCTCCAAACAGAGACCGTGCACATCACTTAATCAGGGC 697
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
    |||
Db 698 ACAAGAAAGAAACAAGTTCAACAAATCATCAAAAGACATTAGGAGATTCAAGAAAGACACT 757
QY 221 LysValAspLysValValIleLeuTyrPheAlaAsnThrGluArgTyrSerLeuVal 240
    |||
Db 758 AAGGTTGACAAAGGTGGTGTCTGTGACAGCAACACAGAGGTATACGCAACGTAGTT 817
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaIleValAspArgAsnGlnAlaGlu 260
    |||
Db 818 GTGGGACTTAAGACACATGGAACCTCTTGCTTCTTGACAGAAATAGGCTGAG 877
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
    |||
Db 878 ATTTCCCTTCAACCTGTATGCAATGCTGTGTATGAAATATGTTCTTTTCAAT 937
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
    |||
Db 938 GGAAGCCCTCAGAACACCTTTGTCCAGACCATATGATCTGCGACATCAAAAGAAATAT 997
QY 301 LeuIleGlyValAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
    |||
Db 998 TTGATTGGAGAGATGACTTTAAGAGTGTCAACCAAAATGAATCGGTGTGGTGGAT 1057
```

QY 321 PheLeuValGlyValGlyLeuValSerProThrSerIleValSerTyraSerHisLeuGlyAsn 340
Db 1058 TTCCTAGTAGGGCCCGCCATCAAGCCATCTATAGTAGTTACAAACATTGGGAAC 1117
QY 341 AasnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleGlySerIleSerIleSer 360
Db 1118 AATGATGCGATGATCTCTCAGCCCTCAAACTTCGCTCGAAGGAATTTCCAAAGAG 1177
QY 361 AasnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrgluProGlyGluHisPro 380
Db 1178 AATGTTGTGACATATGTGTCAACACACACGCACTCTATAGCGCTGTGAACTTCCG 1237
QY 381 AspHisValValIleValSerTyraValProTyraValGlyAspSerIleValAlaMetAsp 400
Db 1238 GACCATGTGCTTCTTAAAGTAGTACCTTACGTTGGGACAGCAAGAGGCGCATGAT 1297
QY 401 GlyTyraThrSerGluIlePheMetGlyGlySerThrIleValLeuHisAsnThrCys 420
Db 1298 GAGTACTCTTGTGAGATATTCAATGGCGGGAAGACACAACTGTGTGACCAACTTGT 1357
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleLeuAlaGlyLeuSer 440
Db 1358 GAGGATTCACTGTACTCTCTATTAATCTTGCACTGTGCTCTTGTGAGCTTACG 1417
QY 441 ThrArgIleGluPheValAlaGluAsnGlyValPheHisSerPheHisProValAla 460
Db 1418 ACTCGAATGAGTTTAAAGCTGAATATGAGGAAGTTCCACTCACTTCCACCAAGTTGG 1477
QY 461 ThrIleLeuSerTyraLeuThrIleValAlaProLeuValProProGlyThrProValAla 480
Db 1478 ACCATCTCAGTTACTGACCAAGGCCCTCTGTTCACCGATACGCGCAAGTGAAT 1537
QY 481 AlaLeuSerIleGluAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1538 GCACTGTCAAGGAGGAGCTATGCTGGAATAACATCTAAGGCGCTGTGTGATTAGCT 1597
QY 501 ProGluAsnAspMetIleLeuGluTyraIleVal 510
Db 1598 CCGAGAACACATGATCTCGAGTACAAAG 1627

RESULT 15
US-09-938-842A-1438
; Sequence 1438, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCLIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1438
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1438

Alignment Scores:
Pred. No.: 3,296-237 Length: 1533
Score: 2378.00 Matches: 448
Percent Similarity: 95.69% Conservative: 40
Best Local Similarity: 87.84% Mismatches: 22
Query Match: 90.35% Indels: 0

DB: 9 Gaps: 0
US-10-718-952-2 (1-510) x US-09-938-842A-1438 (1-1533)
QY 1 MetPheIleGluAsnPheValGlyCysProAsnValIleSerTyraThrGluThrGluIle 20
Db 1 AATGTTATGAGAGGCTTCAAGATGAGAGTCCAAATGTCAMATATACACAGAGATGAGATC 60
QY 21 GlnSerValTyraSerTyrgluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyra 40
Db 61 AATTCGCTGATGATTAACCAACACAGAGTGTCTCCACAGAGATGTAACGCTATCATAT 120
QY 41 GlnThrIleValIlePheProIleSerValIleTyraGluPheValThrAsnIleHisValPro 60
Db 121 CAATGGGTGTGAGAGCAAGACTGTCAATATGATTTCAAAACAGACACTGCTGTCC 180
QY 61 ValLeuGlyValMetLeuValGlyTyraGlyIleAsnAsnGlySerThrLeuThrGlyIle 80
Db 181 AATTTAGGGTTATGCTTGTGTGGGAGGAACAAATGATCAACACTAACCGCTGTGT 240
QY 81 ValIleAlaAsnArgGlyIleSerTyraAlaThrValAspValIleGlnIleAlaAsn 100
Db 241 GTCAATTGTCAACAAAGAAATATCTGGGCACTAAAGACAAAGATGCAACAACTTAAT 300
QY 101 TyraPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 301 TACTTGTGCTTACTTACCTACGCTTATGATTCGTTGCTTCAATAGGTGAAG 360
QY 121 IleTyraAlaProPheIleSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 361 ACTATGCTCTTCAAGAGTCTTCTTCCAAAGTGAATCCAGAGAGATTTGTGTGTGT 420
QY 141 GlyTyraPheIleSerAsnMetLeuAlaAspAlaMetAlaArgAlaValPheAsp 160
Db 421 GGTGGGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 161 IleAspLeuGluIleGluLeuArgProTyraMetGluSerMetLeuProGlyIle 180
Db 481 ATTTGATCTGACAAACACATTAAGCTTACATGAGAGACATATCCACTCTCTGTATC 540
QY 181 TyraAspProAspPheIleAlaAlaAsnGlnGluIleArgAlaAsnAsnValIleValGly 200
Db 541 TATGACCTGATTTATCATGCTCTCTATCAAGGTTACGTCGTCATATGCGTATCAAGGT 600
QY 201 ThrValGlnGluIleValGlnIleIleValAspIleValAlaPheValGluAlaThr 220
Db 601 ACCAAGAAAGAAACAAGTTGATCATCAAGCATCAAGCATGAGGAGTTAAGAGAAAC 660
QY 221 ValValAspValValValIleuThrAlaAsnThrGluArgTyraSerAsnVal 240
Db 661 AAGGTGATTAAGTGTGTGTATGAGCTGCAACACAGACGTTATGCAATGTGAT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaValAlaAspArgAsnGluAlaGlu 260
Db 721 GATGACCTTAACGATACGAGGAGATCTATGCTCTCTGTTGAAGAACAGAGCTGAG 780
QY 261 IleSerProSerThrLeuTyraIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 781 ATCTGCGCTTACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 841 GGAAGTCTCAAAATTAATTTGTCCCGCTTATCGAATGCTATATGCAAGAAATGT 900
QY 301 LeuIleGlyValAspAspPheValSerGlyGlnThrIleMetIleSerValIleValAsp 320
Db 901 TTGATTTGTGGGAGATTTTCAAGAGTGTGATGATGATGATGATGATGATGATGAT 960
QY 321 PheLeuValGlyValGlyIleValProThrSerIleValSerTyraSerHisLeuGlyAsn 340
Db 961 TTCCTTGTGAGAGCTGATTAAGCTTATGATGATGATGATGATGATGATGATGATGAT 1020
QY 341 AasnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIleGlyIleSerIleSer 360

```
Db      1021 AACGACGGGATGAACTCTCAGCGCACGACCTTCCGCTTAAGAGATTCCAAGAGC 1080
Qy      361 AenValValAspAspMetValasnSerAsnAlaIleuTyrGluProGlyGluHisPro 380
Db      1081 AATGTTGTAAGACGACGTGGTAGCTGAATGGCATCTCTTGAGCCTGGTGAACACCT 1140
Qy      381 AspHisValValIleIleTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db      1141 GACCATGTGTGTGTCTCATAGATGTGCCATATGTTGGGATAGTAAAGAGCAATGGAT 1200
Qy      401 GluTyrThrSerGluIlePheMetGlyLysSerThrIleValIleuHisAsnThrCys 420
Db      1201 GAGTACACGCTGACAGATATTCAATGGAGGAGGAGAACACTATAGTGTGCATTAACACTGT 1260
Qy      421 GluAspSerLeuLeuAlaIleProIleIleuAspLeuValIleuLeuAlaGluLeuSer 440
Db      1261 GAGGACTCTCTCTGGCTGCACCAATTATCTTGATCTGTCTCTCGCTGAACCTCAGC 1320
Qy      441 ThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db      1321 ACCAGAAATCCATTTAAGCTGAAGAGAGAGGGGAAGTTCACTCTTCCACCGGTAGCT 1380
Qy      461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db      1381 ACCATTAATCAGTTAACCCTCAAAAGGCACCTTGTACCCCGGAAACACCGGTGTTAAC 1440
Qy      481 AlaLeuSerLysGluArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db      1441 GCTGTGTCGAAGCGCGGGAATGCTGGAATAACATTCTTAGGGCATGTGTGGGCTTGCG 1500
Qy      501 ProGluAsnAsnMetIleLeuGluTyrLys 510
Db      1501 CCAAGAACACATGATCATGAGTACAAG 1530
```

Search completed: June 8, 2005, 05:08:51
Job time : 3728.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 7, 2005, 17:21:22 ; Search time 3731.5 seconds
(without alignments)
5202.406 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632

Sequence: 1 MREBNPKVBCPVKXTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10718952/runat.06062005.173400.12856/app.query.faeta_1.1294
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10718952.@CEN.1.1.6628.@runat.06062005.173400.12856 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:.*
1: gb_esc1:.*
2: gb_esc2:.*
3: gb_hic:.*
4: gb_esc3:.*
5: gb_esc4:.*
6: gb_esc5:.*
7: gb_esc6:.*
8: gb_gsa1:.*
9: gb_gsa2:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2351	89.3	1740	9	CL961092 OaFCC005
2	2064	78.4	2360	1	AJ583520 AJ583520
3	1986.5	75.5	1828	3	CNSG044MK BX827819 Arabidops
4	1597	60.7	1813	3	AK079323 Mus muscu
5	1597	60.7	1814	3	AK005029 Mus muscu
6	1565	59.5	988	7	CK277950 EST724028
7	1452	55.2	936	7	CK279064 EST725142
8	1430	54.3	972	7	CV135801 EST847010
9	1428	54.3	1613	3	CR599696 full-len

10	1420	54.0	953	7	CO414034	CO414034	EST844419
11	1419	53.9	914	7	CK275352	CK275352	EST721430
12	1397	53.1	898	7	CO113870	GR_Eb014	
13	1397	52.2	891	7	CV290142	aoF01-10m	
14	1370	52.1	867	6	CB973407	CAB30002	
15	1367	51.9	865	7	CO085839	GR_Ea03E	
16	1366	51.9	883	6	CD438045	EL01N0508	
17	1359	51.6	906	7	CK320178	CK320178	L2P08a02
18	1357	51.6	912	7	CV142333	EST853542	
19	1353	51.4	892	7	CO409215	CO409215	EST839600
20	1350	51.3	813	7	CK68601	Gm-r1030-	
21	1349	51.3	860	7	CO072025	GR_Ea30P	
22	1338	50.8	842	7	CF211061	CF211061	CAB20007
23	1337	50.8	904	7	CV148446	CV148446	EST859555
24	1333	50.6	891	7	CF212628	CF212628	CGF100065
25	1325	50.3	836	7	CF518404	CF518404	CAP0007_1
26	1316	50.0	860	7	CO082095	GR_Ea46C	
27	1309	49.7	841	7	CO075472	GR_Ea36C	
28	1304	49.5	960	7	CK274458	EST720536	
29	1282	48.7	830	7	CO107630	CO107630	GR_Eb003
30	1276.5	48.5	1466	3	CR600268	CR600268	full-len
31	1276.5	48.5	1513	3	CR606341	full-len	
32	1265	48.1	816	6	CB975215	CAB30006	
33	1265	48.1	831	7	CO132349	CO132349	GR_Eb45C
34	1262.5	48.0	834	7	CK933094	CGF100434	
35	1257	47.8	878	7	CV137015	EST848224	
36	1252	47.6	570	7	CV147272	EST858481	
37	1249	47.5	785	6	CB682609	OSJNEF10C	
38	1229	46.7	800	7	CO112984	GR_Eb004	
39	1229	46.7	843	7	CK267514	EST713592	
40	1228	46.7	795	7	CO113324	CO113324	GR_Eb013
41	1214	46.1	739	4	BG124172	EST7459818	
42	1212	46.0	799	7	CF401745	RTW1_14	
43	1211	46.0	781	6	CB343821	CA32EN000	
44	1210	46.0	778	7	CO071633	GR_Ea30B	
45	1208	45.9	779	6	CD448699	EK07D2305	

ALIGNMENTS

RESULT 1
CL961092 1740 bp DNA linear GSS 21-SRP-2004
DEFINITION OaFCC005463 Oryza sativa Expressed Library Oryza sativa (indica
LOCUS cultivar-group) genomic, genomic survey sequence.
CL961092
CL961092.1 GI:52376905
GSS.

ORIGIN Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Bharatoideae; Oryzaceae; Oryza.
1 (bases 1 to 1740)

REFERENCE Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
Wong,G.K.S., Deng,X.W. and Wang,J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis

JOURNAL COMMENT

Contract: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-8048676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES

source 1..1740
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"

/db xref="taxon:39946"
 /clone lib="Oryza sativa Express library"
 /note="Oryza sativa exon trapped genomic sequences "

ALIGNMENT SCORES:

Pred. No.: 1.45e-272 Length: 1740
 Score: 2351.00 Matches: 449
 Percent Similarity: 94.11% Conservative: 30
 Best Local Similarity: 88.21% Mismatches: 30
 Query Match: 89.32% Indels: 0
 DB: 9 Gaps: 0

US-10-718-952-2 (1-510) x CL961092 (1-1740)

OY 1 Metpneilegluasnphelysvalglucysproasnvallysethrthrluhtgltle 20
 DB 1 AAGTTCATCGAGAGCTTCGCGTGAGAGCCCGACGCGGTACGCGCGCGGAGATC 60
 OY 21 GlnserValIYrAsnTYrGluThrThrgluValHiegluAsnAArgAnslYThrTYr 40
 DB 61 GAGTGGAGACTACAGACACACAGAGAGCTGTGTGCACAGACAGACAGCCGCTCC 120
 OY 41 GlnTPILValIlySProIySserVallySYrGluPhelyThrsnllEhiSValPro 60
 DB 121 CGCTGGGTGTCGCGCCCAAGTCCGTCGCTACAACTTCGAGCACACACCGCTCCC 180
 OY 61 LysleuGlYValMetleuValglYTPrglYglYAsnAnslYSerThrleuThrglYglY 80
 DB 181 AAGCTGGGGGTATGTCGTGGGGGTGGGGGCGGCACACAGGCTCAACGCTGACGGCTGG 240
 OY 81 ValIleAlaAsnAArggluGlyIleSerTPAlaThrlyAspLysleIleGlnIlnAlaAsn 100
 DB 241 GTCATCGCCACAGAGGAGGATCTCATGTGGGACCAAGACAGAGTGCAGCAAGCCAAC 300
 OY 101 TyrPhelYSerleuThrglnAlaSerAlaIleArgValGlySerPheGlnIglYglu 120
 DB 301 TACTATGCTCACTACCCAGGCGTCCACCATCAGGCGTGGAGCTCAACACGGGAGAG 360
 OY 121 IleTYrAlaProPhelySserleuPROMetValAsnProAspAspIleValPheGly 140
 DB 361 ACTTACGCGCCCTTCAGAGGCTCTCGCCCATGTGAACCCCTGATGACCTTGTGTGGG 420
 OY 141 GlyTPAspIleSerAsnMetAsnleuAlaAspAlaMetAlaArgAlaIlyValPheAsp 160
 DB 421 GCGTGGGACATTCAGACACATGAACCTGCGTGCATATACAGGCGCAAGTACTTGAC 480
 OY 161 IleAspLeuGlnLysGlnleuArgProTYrMetGluSerMetleuProleuProglYile 180
 DB 481 ATTGATCTGCAGAGAGAGCTCAGACCTTACATGAGATCATGTGTGCTCTCCCGGCATC 540
 OY 181 TYrAspProAspPheIleAlaAsnGlnIgluArgAlaAsnAsnValIleYsgly 200
 DB 541 TATGACCCCGACTTCATCCGCCCAACAGGGATCCCGGCAACATTCATCAAGGGC 600
 OY 201 ThrlySglnIgluValGlnGlnIleIlelyAspIlelyValAphelySglnAlaThr 220
 DB 601 ACCAAGAAGAGACAGATGAGAGCATCATCAAGACATAGGAGAGTTCAGAGAAAAGAC 660
 OY 221 LysValAspLysValValleuThrAlaAsnThrgluArgTYrSerAnleuVal 240
 DB 661 AAAGTGAAGAGGT 720
 OY 241 ValIglYleuAsnAspThrMetGluAsnleuAlaAlaValAspAArgsnIgluAlu 260
 DB 721 GTTGGGCTCAAGTGAACATGAGAACTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
 OY 261 IleserProSerThrleuTYrAlaIleAlaCYsValMetGluAsnValProheileAsn 280
 DB 781 AATTCACCATTCAGACATGTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
 OY 281 GlySerProGlnAsnThrPheValProglYleuIleAspLeuAlaIleAlaArgsnThr 300

DB 841 GGAAGTCTCTAGAACACCTTTGTGCTGGGCTGATCGATCTTGCTATTAGAACATCTGC 900
 OY 301 LeuIleGlyIYAspAspPheLYSerGlyGlnThrlySmetLYSerValleuValAsp 320
 DB 901 CTGATTTGGTGTATGATTTTCAAGATGTGACAGACAAAGATGAACTGTCTTGTTGAT 960
 OY 321 PheleuValGlyAlaGlyIleYAspProThrsertIleValSerTYrAsnHieGlnYAsn 340
 DB 961 TTCTAGATTGGTGTGAATGAAGCCCACTCAATGTGCACCTCAACACCTTGGGGAAT 1020
 OY 341 AsnAspGlyMetAsnleuSerAlaProGlnThrPheAArgserlySglnIleSerlySer 360
 DB 1021 AATGATGGCATGAACCTTTCGACCTTCACATTCCTCATCAGAGATATCCAAAGGC 1080
 OY 361 AsnValAlaAspAspMetValAsnSerAsnAlaIleleuTYrGluProglYgluHiePro 380
 DB 1081 AATGTGTGATGATGATGATGTCTCAAGCATGCTCTTATGAGACAGACAGAGGCAATGAC 1140
 OY 381 AspHieSValValIleYsTYrValProTYrValGlyAspSerLYsArgAlaMetAsp 400
 DB 1141 GATCATGTTGTGTGATCAAGATATGTCCGTATGTGGAGACAGACAGAGGCAATGAC 1200
 OY 401 GluTYrThrsertGlnIlePheMetGlyGlyLYSerThrIleValleuHieAsnThryS 420
 DB 1201 GAGTACACCTCAGAGATCTTCATGGGGGTAAAGACACATCGTTTGACACACACCTGT 1260
 OY 421 GluAspSerleuLeuAlaAlaProIleIleleuAspLeuValleuAlaGluLeuSer 440
 DB 1261 GAGGACTACCTTCTTGGCGCGCCGATCTTCTTATCTGTGTGTGTGTGTGTGTGTGTGT 1320
 OY 441 ThrArgIleGluPhelysAlaGluAsnGluYlySPhelSserPheHieSProValAla 460
 DB 1321 ACCAGATTTCAGCTGAAGCCGAGGGGAGAGAAATTCATTCCTTCATCCAGTGGCT 1380
 OY 461 ThrIleleuSerTYrleuThrlyAlaProleuValProProglYThrProValValAsn 480
 DB 1381 ACCATCTCGACTACCTTCACAGGACCCCTTGTCTCTGTGACACCGAGTGTGAC 1440
 OY 481 AlaIleuSerLYSglnArgAlaMetleuGluAsnIleMetCArgAlaCYsValIglYleuAla 500
 DB 1441 GCCCTGGCAAGACAGAGGCAATGCTTGAGAAACATCATGAGGCGCTGCGTGGCTGGC 1500
 OY 501 ProGluAsnAsnMetIleleuGluTYr 509
 DB 1501 CCCGAGAACACATGATCTCGAGTAC 1527

RESULT 2
 AJ583520/c 2360 bp mRNA linear EST 08-JAN-2004
 LOCUS
 DEFINITION
 AJ583520 PpProcl Physcomitrella patens cDNA clone pP01015066
 similar to putative myoinositol-1-phosphate synthase, mRNA
 sequence.

ACCESSION
 AJ583520
 VERSION
 AJ583520.1 GI:40781590
 KEYWORDS
 SOURCE
 ORGANISM
 Physcomitrella patens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Funariidae; Funariaceae; Funariaceae; Physcomitrella.
 1 (bases 1 to 2360)
 Hohn,A., Egener,T., Lucht,J.M., Holford,H., Reinhard,C., Schween,G.
 and Reski,R.
 An improved and highly standardised transformation procedure allows
 efficient production of single and multiple targeted gene-knockouts
 in a moss, Physcomitrella patens
 Curr. Genet. 44 (6), 339-347 (2004)

JOURNAL
 COMMENT
 Contact: Schween G
 Plant Biotechnology
 University of Freiburg
 Sonnenstrasse 5, D-70104 Freiburg, Germany.
 Location/Qualifiers
 1..2360
 /organism="Physcomitrella patens"

FEATURES
 source

```

/mol_type="mRNA"
/db_xref="taxon:3218"
/clone="pp001015066"
/tissue_type="protoconema"
/clone_lib="ppprot1"

```

ORIGIN

Alignment Scores:

```

pred. No.:      1,21e-237      Length:      2360
Score:          2064.00        Matches:      396
Percent Similarity: 88.61%      Conservative: 63
Best Local Similarity: 76.45%      Mismatches: 51
Query Match:    78.42%         Indels:      8
DB:             1             Gaps:       3

```

US-10-718-952-2 (1-510) x AJ583520 (1-2360)

```

QY      1 MetPhe11eGluAaPhe1ySValG1uCyProAaVal1ySerThrGluThrGlu1le 20
DB      2150 ATGTTGGTGACAGCTTTCGTGTGAGAGCTCCTCAATGTGAATACACGAGCATGATC 2091
QY      21 GlnSerVal1yTrAaNTyGluThrThrGluLeuVal1leGluAaArg-----Aaang1y 38
DB      2090 AAGTCAGCTATGATGATACCAACACCGAGTTCCTGACAGACGACGCAAGCATGGC 2031
QY      39 ThrTyGlnTriP1eVal1ySPro1ySVal1ySerThrGluThrAaNT1leH1e 58
DB      2030 AAGTATAGAGGGGTACCCCAAGCCCAAGTCCGTCCATACCAATTTCCTCAGCCGCAAG 1971
QY      59 ValPro1ySLeuG1yValMe1leuVal1G1yTrP1yG1yAaAaang1ySerThrLeuThr 78
DB      1970 GTGCCCTAAGCTGGGAATGATGCTCTCGTGGGTGGGAGGACACACGGGTCCACTCTCAG 1911
QY      79 GlyG1yVal1leA1aAaAaArgG1yG1y1leSerTrP1a1aThr1yAaP-Lys1leG1ng1 98
DB      1910 GCGGTATCTCCGCCCAACAAAGAGATCTCATGGGTGACCAAGATGGCGGTGACAGCA 1851
QY      98 n1aAaNTyrrPheG1ySer1eThrGln1a1eSer1a1leArgVal1yG1ySer-PheGln- 117
DB      1850 CCCCAATTACTCTGGATCCCTCACTCAAGCTCTCATGGCGGGGTCTCTTTCAGG 1791
QY      118 --G1yGlu1u1le1yrr1a1aProPhe1ySer1eLeu---PrometValAaP-Propa 135
DB      1790 GTGGAAGAAGATTCATGATCTTTAAAGACATTCACCGATGCTCAACCCCAAC 1731
QY      136 AaP1leVal1PheG1yG1yTrP1a1leSerAaMe1AaAa1aAaP1aMe1a1aArg 155
DB      1730 GATCTCGATCGCGCGGTGGATATCTCAACATGAACTGGCCGATCGCATGATCGT 1671
QY      156 A1a1ySVal1PheAaP1leAaP1leuGln1yGln1leuArgProTyrrMe1G1uSerMe1leu 175
DB      1670 GCCCGGTCTGGACATTCATCTACAGAAAGCACTTCGCTCTCTCATGACGATATGGTA 1611
QY      176 ProLeuProG1y1leTyrrAaPProAaPhe1leA1aAaang1ng1uArg1aAaP 195
DB      1610 CCCATGCTCGGTATCTACACCCCACTTCATCTGCTACCAAGAGCGCTGTGGCAAC 1551
QY      196 AaNTa1le1ySg1yThr1ySg1ng1uGln1yGln1le1le1ySaaP1le1ySaa1a 215
DB      1550 AATTCATCAAGGGTCCAAAGAAAGAGCAATGAGACGATATACAGACATACCGAAC 1491
QY      216 Phe1ySg1uA1eThr1ySVal1aP1ySVal1a1leu1eThrThr1a1aAaThrG1uArg 235
DB      1490 TTCAAGGCCCAACCAAGGTTTACCAAGGTAGTGTGTGACCGCTTCAACAGAGCGT 1431
QY      236 TyrrSerAaNTe1yVal1G1yLeuAaAaPThrMe1-G1uAaNTe1yA1a1a1aVala1a 255
DB      1430 TACAGAGAGATGATCTGGGTCTGAATGACCCCAAGGAATCTTTGATTTCTTGGA 1371
QY      255 pArgAaNTa1a1aG1u1leSerProSerThrLeuTyrra1a1e1a1a1a1a1a1a 275
DB      1370 GAAGGAGAAATCTCGCGCGTGCAGCTTGATGCTTCGCTTGATGCACAGAGCA 1311

```

```

QY      275 nValProPhe1leAaang1ySerProGlnAaThrPheVal1ProG1yLeu1leAaP1eua1 295
DB      1310 GGTTCCTTCATCAACGAGCATCCCAAGAACATTCGCTCTGAGTAATTAATACGC 1251
QY      295 a1leA1aArgAaNT1leu1leG1yG1yAaAaP1e1ySg1yG1nThr1ySMe1y 315
DB      1250 TGTTAAGAAAGACTCTTAATTTGTGTGTGATGACGCTTCAAGAGCGGTCAACGAAGATGA 1191
QY      315 sSerVal1leuVal1aSPhe1e1yG1yA1aG1y1le1yProThrSer1leVal1SerTy 335
DB      1190 ATCCGTTTATGATGATCTTCTAGTTGACGCGGAATTAAGCCCAACGTCATTTGTAGCTA 1131
QY      335 rAaNT1leuG1yAaAaAaP1yMe1e1eNTe1ySer1aProG1ThrPheArgSerTy 355
DB      1130 CAACCATTTGGGCAACAGATGAATGAATCTGTCCGCCCTCAACATTCGTTCCAA 1071
QY      355 eGlu1leSerTySerAaNTa1yA1aAaPMe1e1aAaNTa1le1eNTyrrG1 375
DB      1070 GGAATCTCCAAAGAGCAATGTATGTGACGACATGTGTGCTAGCAATTCATTTATACGC 1011
QY      375 uProG1yGlu1leProAaPhe1yVal1a1le1yTyrrVal1ProTyrrVal1yAaPSe 395
DB      1010 GCCCGGTGAACATCCCGACACGTTATGTGATCAAGTACGTCCCTTATGTGGTGACAG 951
QY      395 r1ySArg1aMe1AaPGLuTyrrThrSerGlu1lePheMe1G1yG1yS1ySerThr1leVa 415
DB      950 CAAGAGGAGCAATGAGAGATGATCACTTCGAGATCTTCATAGGCGGACGTAAACAAATCGT 891
QY      415 l1eNT1aAaNT1ySg1uAaP1eSer1eLeu1aA1aPro1le1le1eAaP1e1yVal1le 435
DB      890 CATGCAATATGCTGGAAGATTTCTTACTGCGCGCGCGCTCATCTTTGACTTGTGCTT 831
QY      435 u1eua1aG1u1e1ySerThrArg1leGluPhe1ySaa1aG1uAaang1y1yS1yPhe11eS1e 455
DB      830 GCTGCGTGAAGCTGTGACAAAGATTCATTAAGAAAGACGACAGAGAAAGTTCATTC 771
QY      455 rPhe11eP1roVal1a1aThr1le1e1ySerTyrrLeuThr1ySaa1aProLeuVal1ProProG1 475
DB      770 TTTCAACCGGTTGCGTCTCTGATCACTCAACAAAGGCTCCCTTGATCCCCCAG 711
QY      475 yThrProVal1a1aAaNTa1le1ySerTyrrArg1a1aMe1e1eGluAaNT1leMe1Arg1a1 495
DB      710 TACCCAGATGTTTAAAGCTTGTGCTTACGACAGAGGCGGATGCTAGAAATATCATGCGCG 651
QY      495 aCyVal1G1yLeuA1aProGluAaAaMe1le1e1yG1yTyrr1yS 510
DB      650 GTGCATTTGGTGTCCCGGACATTAATGATGCTAGAAATCAAG 605

```

RESULT 3
CNS0A4MK
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSTRP424ZE06 of Hormone Treated Callus of strain col-0 of Arabidopsis thaliana (thale cress).
BX827819.1 GI:42462602
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE

CNS0A4MK 1828 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSTRP424ZE06 of Hormone Treated Callus of strain col-0 of Arabidopsis thaliana (thale cress).
BX827819.1 GI:42462602
HTC; GSIUT cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1828)
Castelli,V., Aury,J.M., Jallion,O., Winkler,P., Clepet,C., Menard,M., Cruaud,C., Queller,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Sclanoubat M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
Unpublished
2 (bases 1 to 1828)
Genoscope.
Direct Submission

[illegible][illegible]

Db 302 CTAACTGGCCGACGCGCAGAGTGGCAGAGGACAACTATTATGATCTGTGACCGAC 361
 Qy 108 AAlasera1a1eArgValGly---SerPheGlnGlyGluIleTyrAlaProPheLys 126
 Db 362 GGGGGACCCGCTGAACCTTGGGTCTGGATGAGAACCGCGGAGAGTGTTCCTTCAGT 421
 Qy 127 SerLeuLeuProMetValAsnProAspAlaValPheGlyGlyTyrAspIleSerAsn 146
 Db 422 GGGCTGTACCTCATGTGGTGGCCCAACGACCTGTGTGTGATGGCTGGGATATCTCTGTC 481
 Qy 147 MetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAspIleAspLeuGlnLysGln 166
 Db 482 CTGAACCTGGCCGAGCCCATGGCGCGGCTCTGAGCTCGGCTTGGAGAGAACG 541
 Qy 167 LeuArgProTyrMetGluSerMetLeuProLeuProGlyIleTyrAspProAspPheIle 186
 Db 542 CTGTGGCCCCACATGAGAGCTCGCTCCGCGGCCCTCAGTCTACATTCCTTCAAGTTCATC 601
 Qy 187 AAla1aasngIngluGluArgAlaAsnAsnValIleLysGlyTyrLysGlnGluGlnVal 206
 Db 602 GCTGCCAACACAGACAGACGTCGGGCAACCTCATCCCTGGCAGACAGTCCCAACAGTTG 661
 Qy 207 GlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThrLysValAspLysValVal 226
 Db 662 GAGCAAAATCCGAAGACATTAGAGATTCCGATCCAGTCGCGGAGTTGATTAAGGTCATC 721
 Qy 227 ValLeuTyrThrAlaAsnThrGluArgTyrSerAsnLeuValValGlyLeuAsnAspThr 246
 Db 722 GTGCTGTGACCGCCCAATCGAGAGCCCTTCGCGAGTGTCTCCAGTTCGCAATACACA 781
 Qy 247 MetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlnIleSerProSerThrLeu 266
 Db 782 GCAGAAACTTCTCATCTACTATCCAGCTTGGC---CTGAGAGTGTGCACGCTCCACACTT 838
 Qy 267 TyrAlaIleAlaCysValMetGluAsnValProPheIleAsnGlySerProGlnAsnThr 286
 Db 839 TTGTGCTGTCGACGACTCGAGAGCTCGCGCTTCTCAATGAGATCCCAACAGACACA 898
 Qy 287 PheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeuIleGlyLysAsp 306
 Db 899 CTGGTACCCGCTGCTCCGTAACCTGGCTTCGACGCGCATGTGTTCAGTGTGTATGAC 958
 Qy 307 PheLysSerGlyGlnThrLysMetLysSerValLeuValAspPheLeuValGlyAlaGly 326
 Db 959 TTCAGTCAAGGACAGACTAAGCTCAAGTCTGTCTGTGAGACTTCTCATCGGCTCTGGG 1018
 Qy 327 IleLysProThrSerIleValSerTyrAsnIleSleGlyAsnAsnAspGlyMetAsnLeu 346
 Db 1019 CTCAGAGACCATGTCTCATCTGAGCTATTAACCACTGGGCAACACGCGGACAGAACTTG 1078
 Qy 347 SerAlaProGlnThrPheArgSerLysGlnIleSerLysSerAsnValValAspAspMet 366
 Db 1079 TCTGCACCCGCTGACGTTCCGCTTCAGAGAGGTGACAAAGCAAGTGTGTGACACATG 1138
 Qy 367 ValAsnSerAsnAlaIleLeuTyrGluProGlyGlnIleProAspIleValValIle 386
 Db 1139 GTTCACAGCAACCATGTCTCTACCGGCTGGAGAGCGGCAACACACTGTGGGTGATC 1198
 Qy 387 LysTyrValProTyrValGlyAspSerLysArgAlaMetAspGluTyrThrSerGluIle 406
 Db 1199 AATATGTCTCCCTATGTTCGCGACAGCAACCGGTGCTGGAAGAGTACCTCCGACGCTG 1258
 Qy 407 PheMetGlyGlyLysSerThrIleValLeuHisAsnThrCysGlyLysSerLeuLeuAla 426
 Db 1259 AGCTGTGGGTGGACAAACACTTGGTGTCTCATTAATCTTGGAGGATTCGCTCTCGGCC 1318
 Qy 427 AlaProIleIleLeuAspLeuValLeuLeuAlaGlnLeuSerThrArgIleGluPheLys 446
 Db 1319 GGGCCCATCATGCTGGACCTAAGCTGCTCACAGACTGTGTGACAGGCGCTGAGCTTCTGC 1378
 Qy 447 AlaGluAsnGluGlyLysPheHisSerPheHisProValAlaThrIleLeuSerTyrLeu 466
 Db 1379 ACAGACTCGAGCCCGAGGCTCAGGGCTTCCACACAGTGTCTGCTTGGCTTCTGTC 1438

Qy 467 ThrLysAlaProLeuValProProGlyThrProValValAsnAlaLeuSerLysGlnArg 486
 Db 1439 TTAAAGCCCGGCTTGTGCTCCCGGACGCGCTGTAGATGACCTTCTCCGACAGCGC 1498
 Qy 487 AlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaProGluAsnAsnMetIle 506
 Db 1499 AGCTGTATCGAGAAATATTTTCAGGCGCTTGCCTGGGGCTCCCGCCCAAGACCATGTCTA 1558
 Qy 507 LeuGlyTyrLys 510
 Db 1559 TTAGAGCACAAAG 1570
 RESULT 5
 AK005029
 LOCUS
 DEFINITION
 AK005029 1814 bp mRNA linear HTC 03-APR-2004
 Mus musculus adult male liver cDNA, RIKEN full-length enriched library, clone:1300017C10 product:MHO-INOSITOL 1-PHOSPHATE SYNTHASE A1 (EC 5.5.1.4) homolog [Homo sapiens], full insert sequence.
 ACCESSION
 AK005029
 VERSION
 AK005029.2 GI:26339702
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS
 TITLE
 1 Carrinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komo, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 2 Carrinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komo, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 3 Shibata, K., Itoh, M., Aizawa, K., Nageaka, S., Sasaki, N., Carrinci, P., Komo, H., Akiyama, Y., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sun, N., Ishii, Y., Nakamura, S., Hazama, N., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasaiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 6 (bases 1 to 1814)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arikawa, T., Bono, H., Carrinci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komo, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Oka, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,


```
Qy 367 VALaenSerAsnAlaIleLeuTyrgIuProgiyGluHISproAspHisValValIle 386
Db 1140 GTTCAAGACCAACCATGTCTACGGCGCTGGAGAGCGGCCAGACCATGTGTGATC 1199
Qy 387 LysTyrrValProTyrrValGlyAspSerIysArgAlaMetAspGluTyrrThrsGrtIle 406
Db 1200 AATATATGCTTATGTTGGCCACAGACCGCTTGCTGGACAGACCTCCACAGCTG 1259
Qy 407 PheMetGlyGlyLysSerThrIleValLeuHISAsenThCyGluAspSerLeuAla 426
Db 1260 AGCTGGGGGAGCAACAACCTTGCTGCCATTAACCTGGAGATTCGCTCTCGGCC 1319
Qy 427 AlaProIleIleLeuAspLeuValLeuAlaGluLeuSerThrrArgIleGluPheIys 446
Db 1320 GGGCCATCATGTCTGGACCTAGTGTGCTCACAGAGCTGTCTACAGCGCTGAGCTTCTGC 1379
Qy 447 AlaGluAsnGluGlyLysPheHISserPheHISproValAlaThrIleLeuSerTyrrLeu 466
Db 1380 ACAAGACTCGGACCCGAGCTCAGGGCTTCCACACAGTCTCTCTGTTAGCTTCTCG 1439
Qy 467 ThrIleValProLeuValProProgiTyrrProValIleAsnAlaLeuSerIysGlnArg 486
Db 1440 TTTAAGCCCCCGCTTGCTGCCCGGACCGCTGTATGTAATGCCCTTCCGCCAGGCC 1499
Qy 487 AlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaProGluAsnMetIle 506
Db 1500 AACTGATCGAATAATTTTTCAGGGCTTGGTGGGGCTCCGCCACAGAACACATGCTA 1559
Qy 507 LeuGluTyrrLys 510
Db 1560 TTAGAGCACAG 1571

RESULT 6
LOCUS CK277950 988 bp mRNA linear EST 03-AUG-2004
DEFINITION EST124028 potato abiotic stress cDNA library Solanum tuberosum cDNA
Clone FOAE302 5' end, mRNA sequence.
ACCESSION CK277950
VERSION CK277950.1 GI:39834928
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (bases 1 to 988)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other ESTs: EST724029
Contact: Robin Buell
The Institute for Genomic Research
9713 Medical Center Dr, Rockville, MD 20850, USA
Email: potcdo-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/ .
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers
1..988
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="FOAE302"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TONA"
/clone_lib="potato abiotic stress cDNA library"
/note="Vector: PCWVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
```

following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.:	7,146-178	Length:	988
Score:	1565.00	Matches:	301
Percent Similarity:	96.05%	Conservative:	15
Best Local Similarity:	91.49%	Mismatches:	13
Query Match:	59.46%	Indels:	1
DB:	7	Gaps:	0

US-10-718-952-2 (1-510) x CK277950 (1-988)

```
Qy 55 ThrAsnIleHisValProLysLeuGlyValMetLeuValGlyTyrrGlyLysAsnGly 74
Db 2 ACTGATGTCATGTTCCAAA-TTAGGGGTATAGCTTGTGATGGGGGAGAAACATGTG 60
Qy 75 SerThrLeuThrGlyGlyValIleAlaAsnArgGluGlyIleSerTyrrAlaThrLysAsp 94
Db 61 TCAGCTTGACCTGAGAGGTCTTATTCCTTAACAGAGAAATTCATGCGCTACCAAGAT 120
Qy 95 LysIleGlnGlnAlaAsnTyrrPheGlySerLeuThrGlnAlaSerAlaIleArgValGly 114
Db 121 AAAGTCAGCAAGCCAAATTAATCTTGCTCTTACACAGGCTCTACATTCGAGTTGGT 180
Qy 115 SerPheGlnGlyGluGlnIleTyrrAlaProPheLysSerLeuLeuProMetValAsnPro 134
Db 161 TCTTTCAATGAGAGAGAAATCATGTCCTCTTAAAGACATCTTCCATGGCTCATATCA 240
Qy 135 AspAspIleValPheGlyGlyTyrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAla 154
Db 241 GATGATGATGATCTTGGAGGATGGACATCGACAAATGAATTTGGCAGATGCCATGCA 300
Qy 155 ArgAlaLysValPheAspIleAspLeuGlnLysGlnLeuArgProTyrrMetGluSerMet 174
Db 301 AGGCTTAAGTTTGTGACATTCATCAAAAGCGAGGCTCATCATGGAATCCATG 360
Qy 175 LeuProLeuProGlyIleTyrrAspProAspPheIleAlaAsnGlnGluGlnArgAla 194
Db 361 GCCCAGCGCTGGATCATGATGACCTCATTTATTTGTGCAACCAAGGCTCAGGTGCT 420
Qy 195 AsnAsnValIleLysGlyThrLysGlnGlnValGlnGlnIleIleLysAspIleLys 214
Db 421 AACCATGTCATTTAAAGAAACCAAGAAAGAAATTAATGATCAATATTAAGATTTTACG 480
Qy 215 AlaPheLysGluAlaThrLysValAspLysValAlaValLeuTyrrThrAlaAsnThrGlu 234
Db 481 GAGTTTAAGAGAAACAACAAGATGACAAGGTGGTGGTTTGTGACGTCAACCTGAA 540
Qy 235 ArgTyrrSerAsnLeuValValGlyLeuAsnAspThrMetGluAsnLeuAlaVal 254
Db 541 AGATACACAGAGTGTGCTGTGGCTTATATGATACCAAGAAACCTTTGCTTCTGTG 600
Qy 255 AspArgAsnGluAlaGluIleSerProSerThrLeuTyrrAlaIleAlaCysValMetGlu 274
Db 601 GATAGAAATGAGGCTGAATATCTCTTCCACCTTGATGCTATTGCTTGTATTTCTTGA 660
```


QY 275 AasnValProPhe11easnglySerProGlnAsnThrPheValProGlyLeu1eaePleu 294
 DB 661 AATGCGCCCTTCAATCATGAAAGCCCAAAACACTTTTGTCCAGGCTCATGATTG 720
 QY 295 AAl1ealAaYAsnThrLeu1eeglyGlyAspAspPheLySerGlyGlnThrLyMet 314
 DB 721 GCATCAAGAGAAACACTTGTGATGATGATTAAGAGTGTCAACCAAGATG 780
 QY 315 LySersValLeuValAspPheLeuValGlyAlaGly11eLyProThrSer11eValSer 334
 DB 781 AAGTACAGCTGCTGATTTCTTGTGAGCTGTATTAAGCAACATCAATTTGTAGC 840
 QY 335 TyraAsn11eLeuGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSer 354
 DB 841 TACAACATTTGGGTAAACAATGATGAATGATCTCGGCCCTCAAGACTTCCGCTCA 900
 QY 355 LySGlu11eSerLySerAsnValValAspAspMetValAsnSerAsnAl1eLeuTy 374
 DB 901 AAGGAGATCTCGAAAGCAATGTTGTGATGATGATGATGATGATGATGATGATGAT 960
 QY 375 GluProGlyGlu11eProAspHisVal 383
 DB 961 GAGCTGAGAGACCTGACCATGTT 987

RESULT 7
 CK279064 936 bp mRNA linear EST 03-AUG-2004
 LOCUS EST125142 potato abiotic stress cDNA library Solanum tuberosum cDNA
 DEFINITION clone POAB989 5' end, mRNA sequence.
 CK279064
 ACCESSION CK279064.1 GI:39836042
 VERSION EST
 KEYWORDS
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 936)
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
 Generation of ESTs from abiotic stressed potato tissue
 Unpublished (2003)
 Other ESTs: EST125143
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@igf.org
 Clones can be requested from the University of Arizona Genomics
 Institute via http://genome.arizona.edu/orders/ .
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
 source
 1..936
 Location/Qualifiers
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POAB989"
 /ciseq_type="abiotic stress treated leaf and root tissue"
 /lab_host="DH10B-Tona"
 /clone_11b="potato abiotic stress cDNA library"
 /note="Vector: PCWMSport6.1; Site 1: BcoRI; Site 2: NotI;
 supplier: Solanum tuberosum var. Kennebec plants were
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation
 of the soil with 150 mM NaCl and tissues were harvested at
 following application of the salt stress (leaves: 2hr,
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
 Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering
 applications. Drought stressed plants were harvested after
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
 and 5d). Set 3 were grown under the standard conditions
 and then were cold stressed by placement at 4 C. Cold

stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d. Set 4 were grown under the standard conditions and
 then were heat stressed by placement at 35 C. Heat
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d and 4d and heat-stressed roots were harvested at 6 hr,
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and
 equal RNA from each tissue and stress was pooled to
 construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.:	3,066-164	Length:	936
Score:	1452.00	Matches:	285
Percent Similarity:	95.51%	Conservative:	13
Best Local Similarity:	91.35%	Mismatches:	12
Query Match:	55.17%	Indels:	0
DB:	7	Gaps:	2

US-10-718-952-2 (1-510) x CK279064 (1-936)

QY 163 LeuGlnLySGlnLeuArgProTyrmGluSerMetLeuProGly11eTyraP 182
 DB 3 CTGCAAAAGCAGCTGAGGCCCTACATGAAATCCATGCTTCTTCTGATCTATGAC 62
 QY 183 ProAspPhe11eAla1easnglnGlu1uAga1aAsnAsnVal11eLySGlyThrLyS 202
 DB 63 CCGTACTTCATGACGCTAAACCAAGAGCAGCTGCGCAACAGTATCAAGAACCAAG 122
 QY 203 GlnGlnGlnValGlnGln11eLeuLySerAsp11eLyAlaPheLySGln1aThrLyVal 222
 DB 123 AAAGAAACAGTTGATCAATTTGTTAAAGATTTAAGAGATTCAAGCAAGAAACAGTA 182
 QY 223 AspLyValValValLeuTyrmThra1aAsnThrguArgTySerAsnLeuValGly 242
 DB 183 GACAAGATGATGTTCTATGACGCTCAACCAAGAGGTAAGCAATGTGTTGGC 242
 QY 243 LeuAsnAspThrmGluAsnLeu1a1aValAspArgAsnGlu1aGly11eSer 262
 DB 243 CTTAACGATACCATGAAACCTTTAGCTGCTGAGTGAAGAAAGAGGCTGAATATCT 302
 QY 263 ProSerThrLeuTyra1a1eAlaCyValMetGluAsnValProPhe11easnglySer 282
 DB 303 CTTTCTACATTCATGACGCTATTCCTGTATTCGAAATATGCTTTCTATCAACGAGAC 362
 QY 283 ProGlnAsnThrPheValProGlyLeu1eAspLeuAla1eAlaArgAsnThrLeu1e 302
 DB 363 CCTCAAAACACTTTTGTTCAGAGCTCTTATGATTTGGCCATTAAGAAACCTTTAATT 422
 QY 303 GlyGlyAspAspPheLySerGlyGlnThrLyMetLySerValLeuValAspPheLeu 322
 DB 423 GGTGTGTATGATCTTTAAGAGTGTCAACCAAGAGAAAGTCCGTGTGTATTTCTCT 482
 QY 323 ValGlyAlaGly11eLyProThrSer11eValSerTyraAsn11eLeuGlyAsnAsnAsp 342
 DB 483 GTTGAAGCTGATTAATTAAGCAAGTCAATAGTACCTCAATCACTTGGGTAAACAATGAC 542
 QY 343 GlyMetAsnLeuSerAlaProGlnThrPheArgSerLySGlu11eSerLySerAsnVal 362
 DB 543 GGAATGAATCTTTGCTGCTCAAAACCTTCGGGTAAAGAGATCTCAAAAAGTATGTT 602
 QY 363 ValAspAspMetValAsnSerAsnAla1eLeuTyrmGluProGlyGlu11eProAspHis 382
 DB 603 GTTGTATGACATGTTGCTTGAACCTCAATCTCTTATGATCTGGAGAGACCTTACCAT 662
 QY 383 ValValAla11eLyTyra1aProTyra1aGlyAspSerLyArgAlaMetAspGlyTy 402
 DB 663 GTCTGTGATCAATATGTTCCATATGTTGAGACAGCAAGAGGCAATGATGATGAC 722
 QY 403 ThrSerGlu11ePheMetGlyGlyLySerThr11eValLeuHisAsnThrCySGluAsp 422
 DB 723 ATGTCAAGATTTTTCATGGGCGGAAAGAGCAACATATGTTATGCAACAACCTTGTAGGAC 782

QY 423 SerLeuLeuAlaAlaProIleIleLeuAspLeuValIleuLeuAlaGluLeuSerThrArg 442
Db 783 TCTCTTTGGAGGCTCCAAATTATCTTGACTTGCTTCTTCTGCGAAGTCACTGCG 842
QY 443 IlegluPheLysAlaGlu--AsnGluGlyLysPheHisSerPheHisProValAlaThrI 462
Db 843 ATTCAGCTCAAGCGAAGAGGGAGGGAGAGTTCACCTCTTCACCTGTGTCTACTA 902
QY 462 IeLeuSerThrLeuThrLysAlaProLeuVal 472
Db 903 TTCTCAGCTACTCTCACCAAGGCTCCCTGTGTA 934
RESULT 8
LOCUS CV135801
DEFINITION 972 bp mRNA linear EST 07-SEP-2004
EST847010 Sequencing ESTs from loblolly pine embryos Pinus taeda
CDS clone RP1A131, mRNA sequence.
ACCESSION CV135801 GI:51898116
VERSION CV135801.1 GI:51898116
KEYWORDS EST.
SOURCE Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE Buell, C.R., Hsieh, J., and Cairney, J.
AUTHORS 1 (bases 1 to 972)
TITLE Sequencing of ESTs from loblolly pine embryonic libraries
JOURNAL Unpublished (2004)
COMMENT Contact: C. Robin Buell
Plant Genomics Group
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: rhnell@tigr.org
This clone is available through TIGR. Please contact pine@tigr.org
for further information.
FEATURES
source Location/Qualifiers
1..972
/organism="Pinus taeda"
/mol_type="mRNA"
/cultivar="7-56 mother tree, open-pollinated tree from,
Lyons, Georgia, USA"
/db_xref="taxon:3352"
/clone="RP1A131"
/lab_host="E.coli DH10B-TonA"
/note="Organ: Zygotic Embryo and Megagametophyte. Somatic
Embryo; Vector: pCMV-SPORT 6.1; Site_1: Nott; Site_2:
ScorV; tissue: whole megagametophytes isolated from pine
seeds, whole embryos excised from these megagametophytes,
whole somatic embryos and suspensor tissue from tissue
culture, isolated from cell line A12. Pooled RNA from
zygotic embryos, megagametophytes, and somatic embryos was
used for library construction. Pine cones were harvested
weekly from open-pollinated 7-56 mother trees, collections
occurred from 7/01/02 until 10/15/02. Whole
megagametophytes were first isolated from pine seeds, and
whole embryos excised from these megagametophytes each was
flash frozen. Embryo development was assessed using the
system of Pullman et al (Pullman GS, Johnson S, Peter G,
Cairney J, Xu N. 2003. Loblolly Pine somatic
embryogenesis: development of a maturation medium and
resulting embryo quality. Plant Cell Reports 21:747-758
(http://link.springer.de/link/service/journals/00299/conte
nts/03/00586/). For photographs see Clavatta et al 2001.
Clavatta VT, Morillon R, Pullman GS, Christopels M,
Cairney J. 2001. An aquaglyceroporin is abundantly
expressed early in the development of the suspensor and
the embryo proper of loblolly pine (Pinus taeda L.). Plant
Physiol. 127: 1556-1567
(http://www.plantphysiol.org/cgi/content/full/127/4/1556)

Alignment Scores:

Pred. No.:	1,49e-161	Length:	972
Score:	1430.00	Matches:	282
Percent Similarity:	95.05%	Conservative:	25
Best Local Similarity:	87.31%	Mismatches:	16
Query Match:	54.33%	Indels:	2
DB:	7	Gaps:	0

US-10-718-952-2 (1-510) x CV135801 (1-972)

QY	183	ProSerPheIleAlaAlaAsnGluGluGluValAlaAsnValIleLysGlyThrLys	202
Db	1	CCGATTTTCATAGCGGCAACACCGAGGTAGAGGCAATATATGATCAAGGCTCCAG	60
QY	203	GIngluGluValAlaGluGluIleIleLysAspIleLysAlaPheLysGluAlaThrLysVal	222
Db	61	AAAGAGCAACTGCAAAATATTCAGAGATTCAGAGATTTTAAATGCAAAAGAGTG	120
QY	223	AspLysValAlaValLeuThrThrAlaAsnThrGluArgLysSerAsnLeuValValGly	242
Db	121	GACAAAGTAGTCGTACTGTGACTGTCTAATCCGAAACGATACAGCGATATCTGGGT	180
QY	243	LeuAsnAspThrMetGluAsnLeuLeuAlaValaAspArgAsnGluAlaGluLysSer	262
Db	181	TTAAATGACAGAGAGAAACCCCTGCTCTCTCTGAAAGAAATGAGTCCGAAATCTG	240
QY	263	ProSerThrLeuThrAlaAlaAlaCysValMetGluAsnValProPheIleAsnLysSer	282
Db	241	CCATGACTCTCTACGCGCTTGCGCTTCAGAGAAAGCTTCCTTCATAAAGCGAGT	300
QY	283	ProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeuIle	302
Db	301	CCTCAAAACACCTTGCTTCAGAGATTCATTTAGCAATTCAGAAACAGTATCATC	360
QY	303	GlyGlyAspAspPheLysSerGlyGluThrLysMetLysSerValLeuValAspPheLeu	322
Db	361	GCAGAGAGTACTTCAAGAGCGGCGAGCAAAAGATGAGTGGTACTGGATTTCTTA	420
QY	323	ValGlyAlaGlyIleLysProThrSerIleValSerIleAsnHisLeuGlyAsnAsnAsp	342
Db	421	GTCGATCCCGGCTTAAAGCCAACTTCGATTTGGAGCTACATCACTCGTAAACAGAC	480
QY	343	GlyMetAsnLeuSerAlaProGluThrPheArgSerLysGluIleSerLysSerAsnVal	362
Db	481	GGGATGAATCTGTCAGCTCCCAAGCTTCGCGTCAAGAAATTTCAAGAGCATATGTC	540
QY	363	ValaAspAspMetValAsnSerAsnAlaIleLeuThrGluProGlyGluHisProAspHis	382
Db	541	GTCGACGACATGGTTCAGAGCAAGCCATCTTTACGAAACAGGGGAACACCAAGACAT	600
QY	383	ValValAlaIleLysThrValProThrValGlyAspSerLysArgAlaMetAspGluThr	402
Db	601	GTCGTCGTCATCAATATATGATCTTACGTTGGAGCACAGAGAGCCATGACAGATAC	660
QY	403	ThrSerGluIlePheMetGlyGlyLysSerThrIleValLeuHisAsnThrCysGluAsp	422
Db	661	ACGTCGAGATTTATGAGAGAAACACCTTGTCATCCACAAATACATGTAAAGAT	720
QY	423	SerLeuLeuAlaAlaProIleIleLeuAspLeuValIleuLeuAlaGluLeuSerThrArg	442
Db	721	TGCTTTTTCAGGCTCTCTGATCTTAACCTTGATGCTGCGGCACTCTGACAGG	780
QY	443	IlegluPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAlaThrIle	462
Db	781	ATTCAACTTAAAAAGAAAGC-GAGGGAAATTCATTTCTTCCACCGGTAGCAAGCAT	839
QY	463	IeuSerThrLeuThrLysAlaProLeuValProProGlyThrProValAlaAsnAlaLeu	482
Db	840	TTAAAGTACTCTCAAGAGCGGCACTGGTACCAACAGGACAGCCCTGTGTTAACCCCTG	899
QY	483	SerLysGluArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaProGlu	502
Db	900	GCGAAGCAGAGGCGATCTGGAGAC-ATCATGAGGCGCTGCTGCGCTCCCTCCGAA	958

QY 503 AaapamMet 505
Db 959 AACACATG 967

RESULT 9
LOCUS CR599696
DEFINITION full-length cDNA clone CS0D1009P03 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR599696
VERSION CR599696.1 GI:50480503
KEYWORDS HTC; cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1613)
AUTHORS Li, W.-B., Gruber, C., Jesse, J. and Polayer, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
CONTACT : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue
2 (bases 1 to 1613)
REFERENCE Genoscope.
AUTHORS Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source 1..1613
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1009P03"
/issue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.: 5,72e-161 Length: 1613
Score: 1428.00 Matches: 274
Percent Similarity: 70.55% Conservative: 83
Best Local Similarity: 54.15% Mismatches: 89
Query Match: 54.26% Indels: 60
Gaps: 4

US-10-718-952-2 (1-510) x CR599696 (1-1613)

QY 6 PheUyVAlGlyCyErvOAsnVAllySTyThrGluThGluIleGlnSerValTyAan 25
Db 56 TTCTTCGTCGAAGACCGGAGCGTGTCTACGGCCCGAGGCCATCCAGCGCAATTCGAG 115
QY 26 TyGluThrGluLeuValHnsgIuAanArGAsnGlyThrTyGlnTrpIleValys 45
Db 116 TACCGGACGACCGCGCTCAGCCGCGAG----- 142
QY 46 ProlysSerVallySTyGluPheUyThrAmlIeHValProlysLeuGlyValMet 65
Db 142 ----- 142
QY 66 LeuValGlyTrpGlyCyAynAnsgIySerThrLeuThrGlyCyValIleAlaAnArG 85
Db 143 -----GTTGGCGCTT----- 151
QY 86 GluGlyIleSerTrpAlaThrLyAspLyIleGlnGlnAlaAsnTyPheGlySerLeu 105
Db ----- 105

Db 152 -----CTCAGAGGCGCAACTACGAGCTGCTG 181
QY 106 ThrGlnAlaSerAlaIleArGValGly---SerPheGlnGlyGluGluIleTyAlaPro 124
Db 182 ACTCAGCGCGGACACCTGAGCCCTGAGCCGCGGACCGGAGCGGAGGAGGTTCCGTAACC 241
QY 125 PheUySerLeuLeuProMetValAanProAspPdlIeValPheGlyGlyTrpAplIe 144
Db 242 TTCAGCGCGGTGCTGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 301
QY 145 SerAsnMetAsnLeuAlaAspAlaMetAlaArGAlaIleValPheAspIleAspLeuGln 164
Db 302 TCCTCCCTGACCTGCGCCGAGCGGATCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 361
QY 165 LysGlnLeuArGProTyTrpMetGusSerMetLeuProLeuProGlyIleTyArProAsp 184
Db 362 GAGCACTGTCGCGGACATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 421
QY 185 PheIleAlaAlaAsnGlnGluIleArGAlaAsnValIleLysGlyThrLyAsnGln 204
Db 422 TTCATCGCGGCGCAACGAGCGCGCGGAGACACCTCATCCAGGCTGCGCGGCGGCGG 481
QY 205 GlnValGlnGlnIleIleLysAspIleTyAlaPheLysGlnAlaThrLyAspLyAs 224
Db 482 CAGCTGAGACAGATCCGCGAGGACATCCAGACTTCGCTGAGCGCGGCGGCGGCGGCGG 541
QY 225 ValValIleUyTrpThrAlaAsnThrGluArGTySerAsnLeuValIleLysLeuAsn 244
Db 542 GTCATGTCCTGTCGCGGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 601
QY 245 AspThrMetGluAsnLeuLeuAlaAlaValaAspArGAsnGluAlaGluIleSerProser 264
Db 602 GACACCGCGGAGAACCTGTGCGCACCATTCGCTCGT---CTGAGGAGTGTGCGCTCC 658
QY 265 ThrLeuTyAlaIleAlaCyAsValMetGluAanValProPheIleAsnGlySerProGln 284
Db 659 ACCCTTTCGCGCGCGGCGGACGATCCGAGGCGGTGCTTCCTCAATGGTGTCCGCGAG 718
QY 285 AsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArGAsnThrLeuIleGlyGly 304
Db 719 AACACCGCTGTCGCGGAGACTCTGAGCTCGGTGCGACACCGGCTTTTGTGGCGCGGA 778
QY 305 AspAspPheUySerGlyGlnThrLyAsmElySerValLeuValaAspPheUyValGly 324
Db 779 GATGACTTCAAGTCAGGCGGACCAAGTCAGTCTGTGTGATCTTCCTCATTCGCTTGGC 838
QY 325 AlaGlyIleLysProThrSerIleValSerTyAsnHnIeGluYAsnAnAspGlyMet 344
Db 839 TCCGAGCTTAAGACCATGTCATTCGTAATTACACCACTGGGACAAACATGGGAGAG 898
QY 345 AsnLeuSerAlaProGlnThrPheArGSerLyGluIleSerLySerAsnValaAsp 364
Db 899 AACCTATCGCGCGCATGTCAGTTCGCTCTAAGAGAGGTGCCAAGACGATGGTGGAGC 958
QY 365 AspMetValaAsnSerAsnAlaIleLeuTyGluProGlyGluHnIleProAspHnIleVal 384
Db 959 GACATGTGTCAAGACCAACAGTCTCTATAGCCCGGAGAGAGCTTACCACTGCGGTG 1018
QY 385 ValIleLysTyValProTyValaGlyAspSerLyAspAlaMetAspGlyTyTrpThrSer 404
Db 1019 GTCATCAAGTATGTGCGGTCGAGCTGGGTGACAGCAAGCGGCGGCTGATAGATACCTGG 1078
QY 405 GluIlePheMetGlyLyLysSerThrIleValLeuHnIleAsnThrCyAsGluAspSerLeu 424
Db 1079 GAGCTGATGTCGGGCGGAAACCAACACTGTGTCGACAAACAGCGTGTAGAGACTGCTG 1138
QY 425 LeuAlaAlaProIleIleLeuAspLeuValaLeuAlaGluLeuSerThrArgIleGlu 444
Db 1139 CTGGCGCGACCATATATGTGACCTTACGCTGTGACCGAGCTGTGCGGCGGCGGCGG 1198
QY 445 PheUyAlaGluAsnGluGlyLysPheHnIleSerPheHnIleProValAlaThrIleLeuSer 464
Db 1199 TTCTGCACATGACATGAGACCGGAGCGGAGACCTTCACCCCGTGTCTGCTGCTGCTG 1258

QY 465 TyrLeuThrLysAlaProLeuValProProGlyThrProValValAsnAlaLeuSerLys 484
:::|||||
Db 1259 TTCCTCTTCAAGCGCCCACTACTGCCCCCGGACCGCGTGTGTAATCCGTTTCCGC 1318
QY 485 GlnArgAlaMetLeuGlnAsnIleMetArgAlaCysValGlyLeuAlaProGlnAsn 504
|||||:::|||||
Db 1319 CAGCCAGCTGCATGAGAACATCTCAGGCGCTGCGTGGGCTCCCGCCACAGAACAC 1378
QY 505 MetIleLeuGlnTyrLys 510
|||||:::|||||
Db 1379 ATGCTCTGGAAACACAAA 1396
RESULT 10
CO414034 953 bp mRNA linear EST 02-JUL-2004
LOCUS CO414034
DEFINITION ESF844419 Sequencing ESTs from loblolly pine embryos Pinus taeda
CDNA clone P1M182 5' end, mRNA sequence.
ACCESSION CO414034
VERSION CO414034.1 GI:49630282
KEYWORDS
SOURCE Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE 1 (bases 1 to 953)
AUTHORS Buell,C.R., Heiao,J. and Cairney,J.
TITLE Sequencing of ESTs from loblolly pine embryonic libraries
JOURNAL Unpublished (2004)
COMMENT Contact: C. Robin Buell
Plant Genomics Group
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: rbuell@tigr.org
This clone is available through TIGR. Please contact pine@tigr.org
for further information
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers
1..953
/organism="Pinus taeda"
/mol_type="mRNA"
/cultivar="7-56 mother tree, open-pollinated tree from,
Lyons, Georgia, USA"
/db_xref="taxon:3352"
/clone="P1M182"
/lab_host="E. coli DH10B-TonA"
/clone_11b="Sequencing ESTs from loblolly pine embryos"
/note="Organ: Zygotic Embryo and Megagametophyte, Somatic
Embryo; Vector: pCMV-SPORT 6.1; Site 1: NotI; Site 2:
EcoRV; tissue: Whole megagametophytes isolated from pine
seeds. Whole embryos excised from these megagametophytes,
whole somatic embryos and suspensor tissue from tissue
culture, isolated from cell line A12. Pooled RNA from
zygotic embryos, megagametophytes, and somatic embryos was
used for library construction. Pine cones were harvested
weekly from open-pollinated 7-56 mother trees, collections
occurred from 7/01/02 until 10/15/02. Whole
megagametophytes were first isolated from pine seeds, and
whole embryos excised from these megagametophytes each was
flash frozen. Embryo development was assessed using the
system of Pullman et al (Pullman GS, Johnson S, Peter G,
Cairney J, Xu N. 2003. Loblolly Pine somatic
embryogenesis: development of a maturation medium and
resulting embryo quality. Plant Cell Reports 21:747-758
(http://link.springer.de/link/service/journals/00299/conte
nts/03/00586/). For photographs see Clavatta et al 2001.
(Clavatta VT, Morillon R, Pullman GS, Chiappeels M,
Cairney J. 2001. An aquaglyceroporin is abundantly
expressed early in the development of the suspensor and
the embryo proper of loblolly pine (Pinus taeda L.). Plant
Physiol. 127: 1556-1567
(http://www.plantphysiol.org/cgi/content/full/127/4/1556))
"

ORIGIN

Alignment Scores:

Pred. No.:	2,34e-160	Length:	953
Score:	1420.00	Matches:	273
Percent Similarity:	94.62%	Conservative:	26
Best Local Similarity:	86.39%	Mismatches:	17
Query Match:	53.95%	Indels:	1
DB:	7	Gaps:	0

US-10-718-952-2 (1-510) x CO414034 (1-953)

QY 115 SerPheGlnGlyGlnGlnIleTyrAlaProPheLysSerLeuLeuProMetValAsnPro 134
:::|||||
Db 7 TCCTTTAAGCGGGAAGAGATTTCATGCTCTCTTCAAGACCTCTCTTCCATGCTGATGCT 66
QY 135 AspAspIleValPheGlyGlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetLys 154
:::|||||
Db 67 AATGCAATGCTCTTGGTGGATGGGACATTAACAACATGAATCTAGCGGATGCGCC 126
QY 155 ArgAlaLysValPheAspIleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMet 174
:::|||||
Db 127 CGAGCCAAAGTTCCTTGAATGACCTACAGAAAGCAGCTCGCGCTTACATGCAAGCANTG 186
QY 175 LeuProLeuProGlyIleTyrAspProAspPheIleAlaAsnGlnGlnLysArgAla 194
|||||
Db 187 ACTCCTTCTCGTATATATATGACCCGGATTTCATAGCGGCAACCAAGGTGAGAGGCA 246
QY 195 AsnAsnValIleLysGlyThrLysGlnGlnValGlnGlnIleIleLysAspIleLys 214
|||||
Db 247 AATATATGATCAAAAGGCTCCAAAGAAAGCAACGCAAAAGATATCCAGATATACGA 306
QY 215 AlaPheLysGlnAlaThrLysValAspLysValValIleTyrThrAlaAsnThrGlu 234
|||||
Db 307 GATTTTAATGCAAAACAAAGGTGACAAAGTACGATGCTGATGCTGATATCCGA 366
QY 235 ArgTyrSerAsnLeuValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaVal 254
|||||
Db 367 CGATACAGCAGCATTAATGCTGGGTTTAATGACAGAAAGCAACCTCTGCTCTG 426
QY 255 AspArgAsnGlnAlaGlnIleSerProSerThrLeuTyrAlaIleAlaCysValMetGlu 274
:::|||||
Db 427 GAAACAAATGAGTCCGAAATCTCGCACTCTCTACGCGCTTGCGCTTACAGAG 486
QY 275 AsnValProPheIleAsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeu 294
|||||
Db 487 AACGTTCCCTCATTAAGCGCAGTCTCTCAAAACACCTTCGTCAGATGATGATTTA 546
QY 295 AlaIleAlaArgAsnThrLeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMet 314
:::|||||
Db 547 GCAATTACAAAGAACAGATATCTCCAGAGATGACTTCAGAGCGGCGACAAAGATG 606
QY 315 LysSerValLeuValAspPheLeuValGlyAlaGlyIleLysProThrSerIleValSer 334
|||||
Db 607 AAGTGGGACTCGTGATTTCTTACTGCTGCGGTCTTAAAGCCAACTTCGATGTAAGC 666
QY 335 TyrAsnIleLysGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSer 354
|||||
Db 667 TACATATCCTCGTGAACAGACGAGTGAATCTGTACGCTCCCAACATTTCCGGTCA 726
QY 355 LysGlnIleSerLysSerAsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyr 374
|||||
Db 727 AAGGAATTTTCAAGAGCAATGTCTGCACGACGAGTGTGCGACAAAGCCATCTTTAC 786
QY 375 GluProGlyGlnHisProAspHisAlaValIleLysTyrValProTyrValAlaLysP 394
|||||
Db 787 GAACCAAGGGGAACACCAACCATGTCTGTCAACCAAGTATGATCTTACGTTGAGAC 846
QY 395 SerLysArgAlaMetAspGlnTyrThrSerGlnIlePheMetGlyGlyLysSerThrIle 414
|||||
Db 847 ACCAAGAGACCATGACGATGACGTCGACGATATTTATGGAGGAAAAAC-ACCTT 905
QY 415 ValLeuHisAsnThrCysGlnAspSerLeuAlaAlaProIleIle 430

Db 906 GRATCCACAAATCATGTGAAGATTCGCTTCTTGACGCTCTGTATC 953

RESULT 11

CK275352
LOCUS

DEFINITION CK275352 914 bp mRNA linear EST 03-AUG-2004
EST212430 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POADM85 5' end, mRNA sequence.

ACCESSION CK275352

VERSION CK275352.1 GI:39832330

KEYWORDS EST.

SOURCE

ORGANISM Solanum tuberosum (potato)
Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 914)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

AUTHORS Generation of ESTs from abiotic stressed potato tissue

TITLE Unpublished (2003)

JOURNAL Contact: Robin Buell

COMMENT The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org

Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/

Seq primer: ATT TNG GTG ACA CTA TAG.

Location/Qualifiers

1. 914

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="POADM85"

/tissue_type="abiotic stress treated leaf and root tissue"

/lab_host="DH10B-TonA"

/clone_lib="Potato abiotic stress cDNA library"

/note="Vector: PCWSPort6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).

Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.: 2,89e-160 Length: 914

Score: 1419.00 Matches: 275

Percent Similarity: 97.95% Conservativity: 11

Best Local Similarity: 94.18% Mismatches: 6

Query Match: 53.91% Indels: 0

DB: 7 Gaps: 0

US-10-718-952-2 (1-510) x CK275352 (1-914)
Qy 217 LysGluAlaThrLysValAspLysValValLeuTrpThrAlaIleThrGluArgTyr 236

Db 1 AAGGAGACAGCAAAATGACACAGGTGGTTTGTGACCTGCAACCTGAAGATAC 60

Qy 237 SerAsnLeuValValGlyLeuAsnAspThrMetGluAsnLeuValAlaValAspArg 256

Db 61 AGCAGTGTGCTGTGGCTTGAATGATACAGGAACCTCTTCTCTGTGGATAGA 120

Qy 257 AsnGluAlaGluIleSerProSerThrLeuTyrAlaIleCysValMetGluAsnVal 276

Db 121 AATGAGCCGAAATATCTCTCCACCTTGATGCTTATGCTTGATTTCTTGAATGTC 180

Qy 277 ProPheIleAsnGlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIle 296

Db 181 CCCTTATCATAGGAAGAACCCACAAACATTTTGTGCCAGCCTCATATTTGGCCATC 240

Qy 297 AlaArgAsnThrLeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSer 316

Db 241 AAGAGGAAACATTTGATGTTGGTGTGATGACTTAAAGTGTCAACCAAGATGAAGTCA 300

Qy 317 ValLeuValAspPheLeuValGlyValGlyIleLysProThrSerIleValSerTyrAsn 336

Db 301 GTGCTGTGTTGATTTCTTGTGTGAGCTGGTATTAAGCCAAACATCATTTGAGCTACAC 360

Qy 337 HisLeuGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlu 356

Db 361 CATTTGGTAAACATGATGATGATGATCTGTGGCCCTCAGACTTTCCGCTAAAGAG 420

Qy 357 HisSerLysSerAsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluPro 376

Db 421 ACTCGAAGAAAGCAATGTGTGCATGACATGATGTTTCAATAGCCATCTTATGAGCCT 480

Qy 377 GlyGluHisProAspHisValValIleLysTyrValProTyrValGlyAspSerLys 396

Db 481 GGAAGACACCTGACCATGTTGTGTATGATTAAGTAAACATATGTGAGATACCAAG 540

Qy 397 ArgAlaMetAspGlyTyrThrSerGluIlePheMetGlyGlyLysSerThrIleValLeu 416

Db 541 AAGGCAATGATGATGATGATGATGATTTTCATGTGGTGAAGAAACCATTTGATTC 600

Qy 417 HisAsnThrCysGluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeu 436

Db 601 CACAATACCTGATGAGATCTCTTTGTGGCTGCCAATATCTTGGATTTGGCTCTTG 660

Qy 437 AlaGluLeuSerThrArgIleGluPheLysAlaGluAsnGluGlyLysPheHisSerPhe 456

Db 661 GGTGAATCAGACACCGCATTCAGCTCAAGCTGAAGAGAGAGGATTAATTCACCTCTTC 720

Qy 457 HisProValAlaThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThr 476

Db 721 CATCCGTGGCAGCAATCTCTCACTATCTTACCAAGCTCTCTGTGACCAAGTACA 780

Qy 477 ProValAlaAsnAlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCys 496

Db 781 CCAGTGTGTAATGCCCTTCAAAAGCAGAGGCAAGCTTGAAGAACATATTTAGGGCTTGT 840

Qy 497 ValGlyLeuAlaProGluAsnAsnMetIleLeuGlu 508

Db 841 GTTGAATGTGCAACAGAAACATGATTTGAA 876

RESULT 12

COL113870 998 bp mRNA linear EST 16-JUN-2004

DEFINITION GR_Eb014121.r GR_Eb Gossypium raimondii cDNA clone GR_Eb014121

ACCESSION COL113870

VERSION COL113870.1 GI:48812557

KEYWORDS EST.

SOURCE Gossypium raimondii

ORGANISM Gossypium raimondii

REFERENCE 1 (bases 1 to 998)

AUTHORS

Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Meller, C.,
Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and
Wing, R. A.

TITLE
JOURNAL

Global assembly of cotton ESTs
Unpublished (2004)

COMMENT

Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 014 row: 1 column: 21.

FEATURES

source

Location/Qualifiers

1..898
/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_EB014121"

/tissue_type="floral"

/dev_stage="3 to +3 DEA"

/lab_host="DH10B"

/clone_lib="GR_EB"

/note="Vector: pCWV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:

Pred. No.: 1.29e-157 Length: 898
Score: 1397.00 Matches: 264
Percent Similarity: 95.99% Conservative: 23
Best Local Similarity: 88.29% Mismatches: 12
Query Match: 53.08% Indels: 0
DB: 7 Gaps: 0

US-10-718-952-2 (1-510) x COI13870 (1-898)

49 VALVSTYRGLuPhelysThrAsnIleHisValProlyseuGlyValMetleuValGly 68
1 GTCAAAATGAATTCAGACTGATACCCATGCCCTTAATTTGAGGTGATGCTTGAGGA 60
69 TTPGlyGlyAsnAsnGlySerThrleuThrGlyValIleAlaAsnArgGlyGly 88
61 TGGGAGGAAACAATGCTTCAACCTCACCGGTGGTATTATAGCTAACAAAGGGTATC 120
89 SerTPAlaThrLysAspLysIleGlnIleAlaAsnTyrPheGlySerleuThrGlnAla 108
121 TCTTGGGCTACTACAGACAAGGTACAAAGCGCTAATTACTTGGTTCATTGACTCAAGCA 180
109 SerAlaIleArgValGlySerPheGlnGlyGlyIleTyrAlaProPheLysSerleu 128
181 TCACAGATCCGAATGGGTCTTCAATGAGAAAGATTTATGCTTCATTAAAGGCTT 240
129 LeuProMetValAsnProAspAspIleValPheGlyGlyTyrAspIleSerAsnMetAsn 148
241 CTTCCATGCTGAACCAACCATGATATTGCTTTGAGAGATGGGACATTAGTACATGAAC 300
149 ILeuAlaAspAlaMetAlaArgAlaLysValPheAspIleAspLeuGlnLysGlnLeuArg 168
301 CTAGCTGATGCAATGGCTAGGGCCAAAGCTTTTCGACATGATCTGCAAAAGCAACTGAGA 360
169 ProTyrMetGluSerMetLeuProLeuProGlyIleTyrAspProAspPheIleAlaAla 188
361 CCTTCATGATGAATCCATGGTCCCACTCCCTGGAATTCATGATTCGATTTGATGCT 420
189 AsnGlnGlnGlnArgAlaAsnAsnValIleLysGlyThrLysGlnGlnGlnValGlnGln 208
421 AACCAAGGTGAACGTCCCAATATGCTACCAAGGGGACCAAGAAACAAAGTTCAGCAG 480
209 IleIleLysAspIleLysAlaPheLysGlnAlaThrLysValAspLysValValLeu 228
:::|||||

481 GTCATCAAGAACATCAAGAGCTTCAAGAGAAAACAAAGTGACACAGCTTGTCTACTC 540
229 TTPThraIaAsnThrGlnuArgTyrSerAsnleuValGlyLeuAsnAspThrMetGlu 248
541 TGGACTGCAACATCGAGAGGTACGACATGCTATCGTGGGGCTTAATGACACCGTGGA 600
249 AsnleuLeuAlaAlaValAspArgAsnGlnuAgluIleSerProSerThrleuTyrAla 268
601 AGCTTATGCTCTCTTGGAGAGAAATGATACAGATTTCTCTCCACTTGTATCT 660
269 ILeuAlaCysValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheVal 288
661 ATTGCTTGTCTTGGAAATGTCCTTTCATCAATGAGCGAGCCCAACAAACCTTTGT 720
289 ProGlyLeuIleAspLeuAlaIleAlaArgAsnThrleuIleGlyGlyAspAspPheLys 308
721 CCAGGTTGATTTGATTTGGCTATTCAAGAACTCTGATTTGAGAGATGACTTCAAG 780
309 SerGlyGlnThrLysMetLysSerValleuValAspPheleuValGlyAlaGlyIleLys 328
761 AGTGCCAGACCAAGATGANNATCTGCTCGTGATTTCTTGTGGGGCTGGGATCAAG 840
329 ProThrSerIleValSerTyrAsnHisleuGlyAsnAsnAspGlyMetAsnLeuSer 347
841 CCAACATGATGATGAGATGATACCATCTGAGAAATTAATGATGATCAATCTGTCA 897
RESULT 13
CV290142 891 bp mRNA linear EST 23-SEP-2004
aofo1-10ms3-d10 aofo1 Asparagus officinalis cDNA clone
aofo1-10ms3-d10 5', mRNA sequence.
CV290142
CV290142.1 GI:52575144
EST.
Asparagus officinalis (garden asparagus)
Asparagus officinalis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
Asparagaceae; Asparagus.
1 (bases 1 to 891)
dePamphilis, C., Carlson, J., Ma, H., Soltis, D., Soltis, P.,
Oppenheimer, D., Frohlich, M., Doyle, J., Tanksley, S., Webb, M.,
Leebens-Mack, J., Landherr, U., Ilut, D. and Wall, K.
Generation of ESTs from early male inflorescences of Asparagus
officinalis
Unpublished (2004)
Contact: Claude dePamphilis or James Leebens-Mack
Muller Laboratory
Penn State University
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 863 6413
Fax: 814 865 9131
Email: cw43@psu.edu or jhl10@psu.edu
The sequence provided is trimmed of vector and low quality regions.
Full sequence and original trace file are available from the Plant
Genome Network website (<http://pgn.cornell.edu>)
Plate: aofo1-10ms3 row: d column: 10
Seq primer: M13F
FEATURES
source
1..891
Location/Qualifiers
/organism="Asparagus officinalis"
/mol_type="mRNA"
/db_xref="taxon:4686"
/clone="aofo1-10ms3-d10"
/tissue_type="male inflorescences"
/lab_host="SOLR"
/clone_lib="Aofo1"
/note="Vector: pBluescript SK (+/-); Site 1: EcoRI;
Site 2: XhoI; This is a directionally cloned,
non-normalized library. This library has been generated by
the Floral Genome Project (FGP). The Floral Genome Project
is funded by NSF's Plant Genome Research Program

(DBI-0115684). More information about the project can be obtained at <http://fgrp.bio.psu.edu>

ORIGIN

Alignment Scores:
 Pred. No.: 7,72e-155 Length: 891
 Score: 1374.00 Matches: 268
 Percent Similarity: 97.58% Conservative: 14
 Best Local Similarity: 92.73% Mismatches: 7
 Query Match: 52.20% Indels: 1
 Db: 7 Gaps: 0

US-10-718-952-2 (1-510) x CV290142 (1-891)

```

QY 222 ValAspLeuValValValLeuTTPThRAAsenThrGluabrgfyrSerLeuValVal 241
Db 4 GTRAGAAAGATGATGGTCTTGGAGCTGCAAAATACGAGATGATGCAATGATGATGTC 63
QY 242 G1yleuAsnAspThrMetGluAsnLeuLeuA1a1aValAspArgAsnGluA1a1u1le 261
Db 64 GGTCTCAATGACCAATGAGAACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 123
QY 262 SerProSerThrLeuTyra1a1ealAcYsValMetGluAsnValProphel1eaangly 281
Db 124 TCCCATCAACCTCTATGGTATGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
QY 282 SerProGluAsnThrPheValProGlyLeu1leAspLeuA1a1ealAarGaenThrLeu 301
Db 184 ACCCCACAAACACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
QY 302 1leg1yG1AspAspPheLeysSerG1yInThrLysMetLysSerValLeuValAspPhe 321
Db 244 ATTGGAGAGATGACTTTAAAGATGACACACCAAGATGAAGTCTGTGTTGTTGTTGTTTC 303
QY 322 LeuValG1yA1a1aG1yleysProThrSer1leValSerTyraAsn1sleug1yAsnAsn 341
Db 304 CTGTGGGGGCTGATCAAGCCCACTCATGTGATGATGATGATGATGATGATGATGATGATG 363
QY 342 AspGlyMetAsnLeuSer1a1aProG1InThrPheArgSer1yGlu1leSerLysSerAsn 361
Db 364 GATGGATGAATCTCTGTCACACCAACCTTCCGATCAAAAGAGATGATGAGAACCAAT 423
QY 362 Val1a1aAspAspMetValAsnSerAsnA1a1eLeuTyrg1uProG1yGlu1s1ProAsp 381
Db 424 GTGGTGAATGACATGCTTCAACCAATGGCATCTGTATGAGCTTGCGACACCTTGAT 483
QY 382 H1sVal1a1a1lelystyryValProTyryValG1yAspSerLysArg1a1aMetAspGlu 401
Db 484 CATGTCATGTTATCAAGATATGTCCTGATGTTGATGATGATGATGATGATGATGATGATG 543
QY 402 TyrrThSerGlu1lePheMetG1yG1yLysSerThr1leValLeuH1sAsnThrCyseGlu 421
Db 544 TACACATCTGAGATATTCATGGAGGATGCGACACCATGTGTGTCACAAACCTTGAG 603
QY 422 AspSerLeuLeuA1a1aPro1le1leAspLeuValLeuLeuA1a1aGluLeuSerThr 441
Db 604 GATTCGCTTGGCTGCGCCGATCACTGAGATTTGATTTATGCTGAGCTTACACACC 663
QY 442 Arg1leGluPhe1yA1a1aGluAsnG1yLysPheH1sSerPheH1sProValA1a1aThr 461
Db 664 AGGATTCAGCTCAAGCTGAGGGAGAGGGCAAGTTCCATCTTCATCCAGTGGCTTACC 723
QY 462 1leLeuSerTyrrLeuThrLysA1a1aProLeuValProProG1yThrProVal1a1aAsnA1a 481
Db 724 ATCTTGAGTTACTCTCCCAAGGCTCCCTTGTTCCACAGGSCACACAGGTGGATGCA 783
QY 482 LeuSerLysG1a1aArg1a1aMetLeuGluAsn1leMetArg1a1aCyryValG1yLeuA1a1aPro 501
Db 784 CTCTCAAGCAAGAGGCGATGCTGAGAACATCTCAAGGCTTGACATGGCTTGCTCTCT 843
QY 502 G1uAsnAsnMet1leLeuG1yTyrrLys 510
Db 844 GAGAACAC-ATGATCTCTGAGATCAAA 869

```

RESULT 14

CB973407

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..867

/organism="Vitis vinifera"

/mol_type="mRNA"

/cultivar="Cabernet Sauvignon"

/db_xref="taxon:29760"

/clone="CAB30002_IC_Fc_D10"

/sex="Hermaphrodite"

/dev_stage="Berry stage I"

/lab_host="DH5alpha"

/clone_lib="Cabernet Sauvignon Berry Stage I - CAB3"

/note="Organ: Berry; Vector: pDNR; Site_1: St1; Site_2: St11; CAB3 is a cDNA library of Vitis vinifera 'Cabernet Sauvignon' Clone 8 berries. Samples were collected after berry set from field-grown vines during stage I of berry growth, 17 days after full bloom. The average berry size was 6 millimeters. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGTGTATCAACGAGAGTGGCATTAAGCGCGG-3' and 5'-ATTCTAGAGCGCGGCGCCACATG-dT(30)NN-3'. Library was constructed using the Clontech Creator SMART Kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Alignment Scores:

Pred. No.: 2.25e-154 Length: 867

Score: 1370.00 Matches: 258

Percent Similarity: 95.50% Conservative: 18

Best Local Similarity: 89.27% Mismatches: 13

Query Match: 52.05% Indels: 0

Db: 6 Gaps: 0

US-10-718-952-2 (1-510) x CB973407 (1-867)

```

QY 19 G1uileGlnSerValTyrrAsnTyrrGluThrThrGluLeuValH1sGluAsnArgAsnGly 38
Db 1 GAGATTCAATCTGTGTGACACATATGAGACCAACGAGCTTGTTCATGAGAACAGGATGCG 60
QY 39 ThrTyrrG1nTP1leVal1yAspProLysSerVal1yTyrrGluPheLysThraen1leH1s 58
Db 61 ACCTACCAATGGGTTGTCAAGCCCAAGACTGTCAATATGAATTCATGAGCATGATGCCAT 120
QY 59 ValProLysLeuG1yValMetLeuValG1yTyrrG1yG1yAsnAsnG1ySerThrLeuThr 78

```



```
Db 121 GTCCCATGCTAGGGGATTATGCTTGAGGTTGGGAGGAAACAGCGCTCATCTCTCACT 180
Qy 79 GYGLVVALIIEALAAASARGIUGLYLIESERTPALAThThlyAspLysIleGlnGln 98
Db 181 GGTGGTGTATATGCGCATGCGAGGAGAAATTTATGATGGCAACAGAGACAAAGGTGACCA 240
Qy 99 AIAAENTYrPheGlySerLeuThrglnAlaSerAlaIleArgValGlySerPheGlnGly 118
Db 241 GCGAATTACTTTGGCTCGCTGACTCAGGACATCGACCATCGAGTTGGGCTTCAACATGA 300
Qy 119 GYGLVLIETTYrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleVal 138
Db 301 GAGGAGATTACGCTCCATTCATCAAGAGCCCTCCCTCATGTGTAAACCCAGATGACATTGTT 360
Qy 139 PheGlyGlyTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysVal 158
Db 361 TTGGGGGAGTGGGACATCGATGACATGAACCTGGGACATGATGAGCCAGGCTTAGGTC 420
Qy 159 PheAspIleAspLeuGlnLysGlnLeuArgProTYrMetGlySerMetLeuProLeuPro 178
Db 421 CTGGACATTGATCTGACAGACAGATTGAGGCCCTTCATGAGATCCATGCTCCACTCCCT 480
Qy 179 GYLIETTYrAspProAspPheIleAlaAlaAsnGlnGlnLysArgAlaAsnAsnValIle 198
Db 481 GGAATCTATGACCTGATTTCACTGCTGCCAACCAATGCTCTGTCGCCAACCACTCATC 540
Qy 199 LysGlyThrLysGlnGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlu 218
Db 541 AAAGGACCCAAAGAGAGAGAGGTCCAGCAGATCATTTAAAGATTTAAAGGTTTAAAGG 600
Qy 219 ALAThLysValAspLysValValLysLeuTYrThrAlaAsnThrglnArgTYrSerAsn 238
Db 601 GCAACCAAGGTGGATAGGTGAGTGGTGGCTGTGAGACTGCCAACAGAGGATACAGTAT 660
Qy 239 LeuValValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGlu 258
Db 661 TTGATGTTGGGACATGATACACATAGAAATCTCTGGCTTCTGATGATGATGATG 720
Qy 259 ALAGLIIESeProSerThrLeuTYrAlaIleAlaCysValMetGluAsnValProPhe 278
Db 721 CCTGAGATCTCTCTTCCACCTTGATGCTGTGGCTGTGTCATGAGAAATGTTCCCTTT 780
Qy 279 ILeAsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArg 298
Db 781 ATCATATGAGNGCCACAGAAACATTTTGTTCAGAGGCTGATTGATTGGCTATTAAAGG 840
Qy 299 AsnThrLeuIleGlyGlyAspAspPhe 307
Db 841 AACAGTTGATGTGGTGGGATGATTTT 867

RESULT 15
CO085839 865 bp mRNA linear EST 16-JUN-2004
LOCUS GR_Ea03E19.f GR_Ea Gossypium raimondii cDNA clone GR_Ea03E19 5',
DEFINITION mRNA sequence.
ACCESSION CO085839
VERSION CO085839.1 GI:48776473
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 865)
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
COMMENT
TITLE
JOURNAL
CONTACT: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
```

```
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 03 row: E column: 19.
location/Qualifiers
1. 865
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea03E19"
/issue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_id="GR_Ea"
/notes="vector: pCMVSPORT-6.1; Site 1: NotI; Site 2:
EcoRV; library made by Invitrogen with RNA supplied by
Wendle lab. Directional clones into NotI-EV. Clones
plated/picked by Agt. More glycerol clones held in -80."
```

ORIGIN

Alignment Scores:

Pred. No.:	5,18e-154	Length:	865
Score:	1367.00	Matches:	262
Percent Similarity:	97.22%	Conservative:	18
Best Local Similarity:	90.97%	Mismatches:	8
Query Match:	51.94%	Indels:	0
DB:	7	Gaps:	0

US-10-718-952-2 (1-510) x CO085839 (1-865)

```
Qy 143 AspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAspIleAsp 162
Db 2 GACATTATGACATGAACCTAGCTGATCTCTATGCTGAGCCCAAGTTTTCGACATGAT 61
Qy 163 LeuGlnLysGlnLeuArgProTYrMetGlySerMetLeuProLeuProGlyTYrIleTYrAsp 182
Db 62 CTGCAAAAGCACTGAGACCTTACATGGAATCCATGATGCCATCTCTGGAATCTACAT 121
Qy 183 ProAspPheIleAlaAlaAsnGlnGlnLysArgAlaAsnValIleLysGlyThrLys 202
Db 122 CCTGATTTCACTTGTCTGCTACCAAGGTGAACCTGCCAATATGTCATCAAGGACCAAG 181
Qy 203 GlnGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnValThrLysVal 222
Db 182 AAAGAAACAAGTTCAAGCAGTCAATCAAGACATCAAGAGATTCAAGAGAAACAAAGGTG 241
Qy 223 AspLysValValLysLeuTYrThrAlaAsnThrglnArgTYrSerAsnLeuValValGly 242
Db 242 GACAAAGTTGTTGTACTGTGACCTGCAACACCTGAGAGGTACAGCAATGTCATCTGGGG 301
Qy 243 LeuAsnAspThrMetGluAsnLeuLeuAlaValAspAsnGlnValAGLIIESePro 262
Db 302 CTAAATGACACCGGTGAAAGCCTTAGCTCTTTGGAGAAATGATCAGAGATTCT 361
Qy 263 ProSerThrLeuTYrAlaIleAlaCysValMetGluAsnValProPheIleAsnGlySer 282
Db 362 CTTCCATTGTGTATGCTATGCTGTGTTCTTGAAATGTTCTTTCATCATGAGCAGC 421
Qy 283 ProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeuIle 302
Db 422 CCACAAACACCTTTGTTCCAGGGTGTGATTGCTATTCAAAGAACCTGTGATT 481
Qy 303 GYGLYAspAspPheLysSerGlyGlnThrLysMetLysSerValLysValAspPheLeu 322
Db 482 GAGGAGATGACTTAAAGATGAGCCAGACCAAGATGAAATCTGCTCTGTGATTTCTT 541
Qy 323 ValGlyAlaGlyIleLysProThrSerIleValSerTYrAsnHisLysGlyAsnAsnAsp 342
Db 542 GTTGGGCTGGGATCAAGCCAAATCGATGATGAGTTCAACCAATCTGGAAATATGAT 601
Qy 343 GlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlyLysSerLysSerAsnVal 362
```


Db	662	GGCATGATCTGTGACGACCCCAACCTTCGGTCCAAAGAGATCTCAAGAGCAATGTT	661
Qy	363	ValAspAspMetValAsnSerAsnAlaIleLeuTYrGluProGlyGluHisProAspHis	382
Db	662	GTGTGTGACATGGTTTCAAGCAATGGAAATCCGTATGACCTGGTGAAACATCCTGATCAT	721
Qy	383	ValValIleIleValSTyTyrValProTYrValGlyAspSerIleValGluIleAspGluTYr	402
Db	722	GTGTGGTCATCAAGTATATGTGCCATATGTGGGAGACGACAGAGACCATGATGAGTAC	781
Qy	403	ThrSerGluIlePheMetGlyGlyIleSerThrIleValLeuHisAsnThrCysGluAsp	422
Db	782	ACATCAGAGATATTATCATGGAGGACAAGAACACATTGTGTGACAACAACATGTAGAGAT	841
Qy	423	SerLeuLeuAlaIleProIleIle	430
Db	842	TCCCTGTGGCTGCTCCCATTAATC	865

Search completed: June 7, 2005, 22:49:15
Job time : 3757.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 7, 2005, 17:11:57 ; Search time 5315 Seconds
(without alignments)
4649.513 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631

Sequence: 1 MFENFKVSEPVKXTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPO1.spool.p/US10718952/runat_06062005.173400.12846/app_query.fasta_1.1294
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOCL=0 -LOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10718952@cgn2.1_1.7010@runat_06062005.173400.12846 -NCPU=6 -ICPU=3
-NO MAP -LARGEBUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_da:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pal:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_scs:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2593	98.6	1739	8	AY038802 Glycine m
2	2593	98.6	1782	6	BD075266 Soybean P
3	2588	98.4	1533	6	BD075269 Soybean P
4	2580	98.1	1791	8	AF293970 Glycine m

5	2475	94.1	1959	8	AB009881	AB009881 Nicotiana
6	2470	93.9	1931	6	AB032073	AB032073 Nicotiana
7	2470	93.9	1950	6	E27176	E27176 Novel INPS
8	2468	93.8	1959	8	BT013505	BT013505 Lycopersi
9	2452	93.2	1845	8	AF284065	AF284065 Sesamum 1
10	2437	92.6	1954	8	BT013759	BT013759 Lycopersi
11	2400	91.2	1538	8	AF120146	AF120146 Trifolium
12	2400	91.2	1538	8	AF120147	AF120147 Trifolium
13	2400	91.2	1538	8	AF120148	AF120148 Trifolium
14	2399	91.2	1866	8	AF433879	AF433879 Suaeda ma
15	2395	91.0	2053	8	MCU32511	U32511 Mesembryant
16	2384	90.6	1533	6	CQ080508	CQ080508 Sequence
17	2384	90.6	1533	6	AX056743	AX056743 Sequence
18	2384	90.6	1533	8	AY143904	AY143904 Arabidops
19	2384	90.6	1663	8	AY054202	AY054202 Arabidops
20	2384	90.6	1664	8	AY053415	AY053415 Arabidops
21	2376	90.3	1564	8	AY096554	AY096554 Arabidops
22	2376	90.3	1781	6	AX054630	AX054630 Sequence
23	2376	90.3	1781	8	BNT66307	BNT66307 Brassica na
24	2376	90.3	1863	8	AY065415	AY065415 Arabidops
25	2373.5	90.2	1845	8	AY028259	AY028259 Arabidops
26	2361	89.7	1533	8	AY323824	AY323824 Xerophyta
27	2361	89.7	1890	8	ATU30250	ATU30250 Arabidops
28	2358	89.6	1842	8	SPM1PHSYM	SPM1PHSYM
29	2354	89.5	1535	6	BD073472	BD073472 Regulated
30	2354	89.5	1665	6	BD073470	BD073470 Regulated
31	2354	89.5	1665	8	AF056326	AF056326 Zea mays
32	2354	89.5	1936	8	AB059557	AB059557 Avena sat
33	2352	89.4	1914	8	AK103501	AK103501 Oryza sat
34	2352	89.4	1915	8	AK058750	AK058750 Oryza sat
35	2349.5	89.3	1567	8	BT001931	BT001931 Arabidops
36	2349.5	89.3	1837	8	AY085931	AY085931 Arabidops
37	2349.5	89.3	1870	8	AF372954	AF372954 Arabidops
38	2336	89.2	1931	6	AR137808	AR137808 Sequence
39	2336	89.2	1931	6	AR170167	AR170167 Sequence
40	2345	89.1	1868	8	AB012107	AB012107 Oryza sat
41	2338	88.9	2152	8	AF056325	AF056325 Hordeum v
42	2337	88.8	1959	6	AX151207	AX151207 Sequence
43	2335.5	88.8	1921	8	AT004876	AT004876 Arabidops
44	2328.5	88.5	1871	8	PV038920	UV38920 Phaseolus v
45	2327.5	88.5	1978	8	CPINOIG	232632 C.paradisi

ALIGNMENTS

RESULT 1	AY038802	1739 bp	mRNA	linear	PLN 16-JUL-2001
LOCUS	AY038802				
DEFINITION	Glycine max myo-inositol-1-phosphate synthase mRNA, complete cds.				
ACCESSION	AY038802				
VERSION	AY038802.1	GI:14764465			
KEYWORDS					
SOURCE					
ORGANISM	Glycine max (soybean)				
REFERENCE	1 (bases 1 to 1739)				
AUTHORS	Hitz,W.D., Carlson,T.J., Kerr,P. and Sebastian,S.				
TITLE	Biochemical and Molecular Characterization of a Mutant that				
JOURNAL	Confers a Decreased Raffinosecarbitide and Phytic Acid Phenotype on				
REFERENCE	2 (bases 1 to 1739)				
AUTHORS	Carlson,T.J. and Hitz,W.D.				
TITLE	Direct Submission				
JOURNAL	Submitted (06-JUN-2001) Crop Genetics, DuPont Co., P.O.Box 80402,				
FEATURES	Wilmington, DE 19880-0402, USA				
SOURCE	Location/Qualifiers				
	1..1739				
	/organism="Glycine max"				
	/mol_type="mRNA"				

CDS
/cultivar="Wye"
/db_xref="taxon:3847"
/dev_stage="developing embryo"
11..1543
/EC_number="5.5.1.4"
/function="cyclizes glucose 6-phosphate to
L-myo-inositol-1-phosphate"
/note="mIPS-1A"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAK72098.1"
/db_xref="GI:14764466"

/translation="MEIENFKVCEPNVXETETEQSYNYNETTELHENRNGYOMIV
KPSYKYEFTKNIHVPKLGVMVLVGWGNNGSTLGGVLANREGISMATKDKIQANF
GSLTOASAIRVGSFOGEBIYAPFKSLIPVNPDDIVPGMDISNNLADAMARAFVD
IDLOKOLAPYMSMLPLGIYDPDRIANOBERANNVKGTCBOVOOIIIDIKAFKE
ATKDVAVVLTANTERYSNLVGLNDTMENTLAAVDNRNEISPSIIPALACVENY
PFINGSPOFTPEVPLIDLAIAENTLIGDDFKSGQTKMSVLDPLVGAIGKPSIVS
YNHLGNNDGMNLSAPQTFRSKEISKSNVDDVNSNALLYEPGEPHDVVVLYKVPY
GSKRAMDEYTSIEFMGKSTVLNTECDLSLAAPILIDVLIAELSTRIEFKANE
GKEHSHFHVATILSYLTKAPLIVPGTPVNVNLSKQRAMLENIMRACVGLAPENNMLE
YK"

ORIGIN

Alignment Scores:
Pred. No.: 2,92e-199 Length: 1739
Score: 2593.00 Matches: 502
Percent Similarity: 99.22% Conservative: 4
Best Local Similarity: 98.43% Mismatches: 4
Query Match: 98.56% Indels: 0
Gaps: 0

US-10-718-952-12 (1-510) x AYO38802 (1-1739)

QY 1 MetPhe11eGluAsnPhelysValG1uSerProAsnVal1yETrThrG1uThrG1u1le 20
DB 11 ATGTTTCATCGAAGATTTTAAAGTTGAGTGTCTTAATGTGAGTACACCAGACTGAGATT 70
QY 21 G1nSerVal1TyAsnTrG1uThrThrG1uLeuValH1eG1uAsnAryGsnG1yThrTYr 40
DB 71 CAGTCGGTGTACACACGAAACACGAACTGTGTACGAGAAACGGAATGACACTAT 130
QY 41 G1nTrP1eVal1ySerPro1ySerVal1AsnTrG1nPhelysThrAsnThrH1sValPro 60
DB 131 CAGTGGATGTCGAACCCAAATCTGTCAATACGAATTTAAAAACCAACATTCATGTTCT 190
QY 61 LysLeuG1yValMetLeuValG1yTrP1yG1yAsnAsnG1ySerThrLeuThrG1yG1y 80
DB 191 AAATTAGGGGTATGCTGTGGGTGGGTGGAAACACGCGCTCAACCTCTCACCGTGGT 250
QY 81 Val11eAlaAsnAryG1uAsp11eSerTrP1a1ThrLysAspLys11eG1nG1nAlaAsn 100
DB 251 GTTATTGCTTAACGAGGGGCAATTCATGAGGCTACAAAGACAAATTTCAACAAAGCCAT 310
QY 101 TyrPheG1ySerLeuThrG1nAlaSerAla11eAryValG1ySerPheG1nG1yG1u 120
DB 311 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGCTTCAGGAGAGGAA 370
QY 121 11eTyra1PProPhe1ySerLeuLeuPProMetVal1AsnProAsp11eVal1PheG1y 140
DB 371 ATCTATGCCCAATTCAGAGCTGCTTCCAAATGTGTTAACCTTGACACCACTTGTTGGTGG 430
QY 141 G1yTrPAsp11eSerAsnMetAsnLeuAlaAspAlaMetAlaAryAla1yVal1PheAsp 160
DB 431 GGATGGGATATACGACAACTGAACCTGGCTGATGCTCAAGCCAGGCAAGG1GTTGAC 490
QY 161 11eAspLeuG1nLysG1nLeuAryPro1yMetG1uSerMetVal1ProLeuProG1y1le 180
DB 491 ATCGATTTCAGAGAGAGTGAAGCCTTTCATGAGATTCATGCTTCCATCCCGGAATC 550
QY 181 TyraAspProAspPhe11eAlaAlaAsnG1nG1uAryAlaAsnAsnVal11eLysG1y 200
DB 551 TATGACCCGGAATTTCTGCTGTCCAAACGAGAGAGCGTGCACAAACGCTCATCAAGGAC 610

QY 201 Thr1ySG1nG1uG1nVal1G1nG1n11e11eLysAsp11eLysAlaPheLysG1uAlaThr 220
DB 611 ACAAAGCAAGGACGAAAGTTCAACAATATCTCAAGACATCAAGGGGTTTAAAGAGCCACC 670
QY 221 LysVal1AspLysVal1Val1LeuTrPThrAlaAsnThrG1uAryG1ySerAsnLeuVal 240
DB 671 AAAGTGCAAGAGTGGTGTACTGTGGACTGCGCAACAGAGAGGTACAGTAATTTGGTT 730
QY 241 ValG1yLeuAsnAryPThrMetG1uAsnLeuLeuAla1Val1AspAryAsnG1uAlaG1u 260
DB 731 GTGGGCTTAAATGACACATGAGATCTTGGCTGCTGTGGACAGAAATGAGGCTTAG 790
QY 261 11eSerProSerThrLeuTyra11eAlaCyValMetG1uAsnVal1ProPhe11eAsn 280
DB 791 ATTCTCTTCCACTTGTATGCCATGCTGTGTATGAGAAATGTTCTTTCTATTAT 850
QY 281 G1ySerProG1nAsnThrPheVal1ProG1yLeu11eAspLeuAla11eAlaAryAsnThr 300
DB 851 GGAAGCCCTCAGAACACTTTGTACCAAGGCTGATTCCTTGGCCATCGCAGAGAACCT 910
QY 301 Leu11eG1yG1yAspAspPheLysSerG1yG1nThrLysMetLysSerVal1LeuVal1Asp 320
DB 911 TTGATTGGTGGAGATGACTTCAAGAGTGTCAAGCAAAATGAATCTGTGTGGTAT 970
QY 321 PheLeuVal1G1yAlaG1y11eLysProThrSer11eVal1SerTyraSnH1eLeuG1yAsn 340
DB 971 TTCTCTTGGGGGCGTGTATCAAGCCAACTATATGCTTTCACCAACATCTGGGAAC 1030
QY 341 AsnAspG1yMetAsnLeuSerAlaProG1nThrPheArySerLysG1u11eSerLysSer 360
DB 1031 AATGATGTGATGAATCTTTCGCTCCCAAACTTTCCTCCCAAGGAATCTCCCAAGAGC 1090
QY 361 AsnVal1Val1AspAspMetVal1AsnSerAsnAla11eLeuTyrg1uProG1yG1uH1sPro 380
DB 1091 AACGTTGTGATGATATGTCACAGCAATGCAATCTCTATAGAGCTGTGTAACATCA 1150
QY 381 AspH1sVal1Val11eLysTyra1ProTyraVal1G1yAspSerLysAryAlaMetAsp 400
DB 1151 GACCATGTTGTTTATTAAGTATGCTCTTACGTAGGGGACAGCAAGAGCCATGAT 1210
QY 401 G1uTyTrSerG1u11ePheMetG1yG1yLysAsnThr11eVal1LeuH1sAsnThrCyS 420
DB 1211 GAGTACACTTCAGAGATATTCATGAGTGGTGAAGAGCACTTGTTCACACACATGC 1270
QY 421 G1uAspSerLeuLeuAlaPro11e11eLeuAspLeuVal11eLeuAlaG1uLeuSer 440
DB 1271 GAGGATCCCTCTTAAGCTCTCTTAATCTTGAATTTGCTGCTGAGGCTCAGC 1330
QY 441 ThrATG11eG1nPhelysAlaG1uAsnG1uG1yLysPheH1sSerPheH1sProVal1Ala 460
DB 1331 ACTAATAATCAGATTAAAGCTGAATAATGAGGAAATTCACATCTTCCACCCAGTTGCT 1390
QY 461 Thr11eLeuSerTyraLeuThrLysAlaProLeuVal1ProProG1yThrProVal1Val1sn 480
DB 1391 ACCATCTTCAGCTACCTCAACCAAGGCTCTGCTTCCACCGGGTACCAAGTGTGAT 1450
QY 481 AlaLeuSerLysG1nAryAlaMetLeuG1uAsn11eMetAryAlaCyVal1G1yLeuAla 500
DB 1451 GCATTGTAAAGACAGCGGCAATGCTGAAAAACATAATGAGGGCTTGTGTGATGGCC 1510
QY 501 ProG1uAsnAsnMet11eLeuG1uTyTrLys 510
DB 1511 CCAAGAAATACATGATCTCGAGTACAAAG 1540

RESULT 2
LOCUS BD075266 1782 bp DNA linear PAT 27-AUG-2002
DEFINITION Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid.
ACCESSION BD075266
VERSION BD075266.1 GI:22620869
KEYWORDS JP 2001519665-A/1.

SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1782)
AUTHORS Hitz,W.D. and Sebastian,S.A.
TITLE Soybean plant producing seeds with reduced levels of raffinose
JOURNAL Patent: JP 2001519665-A 1 23-OCT-2001;
BI DU PONT DE NEMOURS AND CO
COMMENT OS Soybean line LR13
PN JP 2001519665-A/1
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PI 08-APR-1997 US 08/835751
PR WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN
PC C12N15/52,C12N15/82,C12N15/11,C12N9/90,A01H5/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Soybean plant producing seeds with reduced
levels of raffinose
CC saccharides
CC and phytic acid
FH Key Location/Qualifiers
FT CDS 54..1586.
Location/Qualifiers
1..1782
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:

Pred. No.:	3e-199	Length:	1782
Score:	2593..00	Matches:	502
Percent Similarity:	99.22%	Conservative:	4
Best Local Similarity:	98.43%	Mismatches:	4
Query Match:	98.56%	Indels:	0
DB:	6	Gaps:	0

US-10-718-952-12 (1-510) x BD075266 (1-1782)

```

QY      1 MetpheiLegluAsnPhelysValGluSerProasnValysTyThrGluThrGluIle 20
DB      54 ATGTTTCATGAGAAATTTTAAGTTGAGTCTCTTAAGTGAAGTACCCAGACTGAGATT 113
QY      21 GluSerValTyraSnTyGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTy 40
DB      114 CAGTCCGTGTACAACTACGAAACCAACCGAATTGTTCAAGAGAACGAAATGCGACTAT 173
QY      41 GINTPILlValLysProLysSerValAsnTyrgInPhelyThrAsnThriSValPro 60
DB      174 CAGTGAATGTCAAAACCCAAATCTGTCAAAATCGAAATTTAAACCAACATTCATCTTCT 233
QY      61 LysLeuGlyValMetLeuValGlyTPGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB      234 AAATTAAGGGGTAAATGCTTGTGGGTGGGTGGGTGAACACACGCTCAACCTTCACCGGTGT 293
QY      81 ValIleAlaAsnArgGluAspIleSerTPAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB      294 GTTATTGCTTAACCGAGAGGCAATTCATGCGCTCAACAGACAAAGATTCAACAGCCAAAT 353
QY      101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGlu 120
DB      354 TACTTTGGTCTCCCTCAACCCCAAGCTTCAGTATCCGAGTGGGTCTTCCAGGAGAGGAA 413
QY      121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB      414 ATCTATGCCCATTCAGAGGCTGCTCTCAATGGTTAACCCTGACGACATTTGTGTTGGG 473
QY      141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB      474 GGATGGAGTATATCAGCAACATGAACCTGCTGATGTCATGCGCAGGAGCAAGGTGTTGAC 533
  
```

```

QY      161 IleAspLeuGlnLysGlnLeuAspProTyMetGluSerMetValProLeuProGlyIle 180
DB      534 ATCGATTTCAGAGACAGTTGAGGCTTTCATAGAAATCATGCTTCACTCCCGGAATC 593
QY      181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB      594 TATGACCCGATTTATCTGCTGCTCAACCAAGAGAGCGTGCCAAACAACTCATCAAGGCG 653
QY      201 ThrLysGlnGlnGlnValGlnGlnIleIleLysAspIleLysValPheLysGluAlaThr 220
DB      654 ACAAAAGCAAGACCAAGTTTCACAAATCATCAAAAGATCAAGCGCTTTAAGAGAGCCACC 713
QY      221 LysValAspLysValValLeuTPThrAlaAsnThrGluAspTyrSerAsnLeuVal 240
DB      714 AAAGTGGACAAAGGTGTTGTACTGTGAGTGCACACAGAGAGTACGTAATTTGGTT 773
QY      241 ValGlyLeuAsnAspThrMetGluAsnLeuValAlaValAspArgAsnGluAlaGlu 260
DB      774 GTGGGCTTAATAGACCAATGAGAAATCTTGGCTGCTGTGAGACAGAAATGAGCTGAG 833
QY      261 IleserProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB      834 ATTTCTCTTCCACCTTGATGCCATTCCTGTGTATGAAATGTTCTTTCATTAAT 893
QY      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB      894 GGAAGCCCTCAGAACACTTTGTACCAAGGCTGATGATGATCTTCCATCCGAGAACACT 953
QY      301 LeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB      954 TTGATTGGTGAGAGTACCTTCAAGAGTGTCTAGACCAAAATGAATCTGTGTGTTGAT 1013
QY      321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB      1014 TTCCTTGGGGGCTGTATATCAAGCCAACTATATAGTACATCAACCATCTGGGAAAC 1073
QY      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB      1074 AATGATGGTATGAATCTTGGCTCCACAACTTTCCTCCAAAGAAATCTCCAAAGAC 1133
QY      361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB      1134 AACGTTGTTGAGATATGCTCAACGACATGCAATCCTCTATAGACCTGTGAATCTCA 1193
QY      381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB      1194 GACCATGTTGTATTATTAAGTATGCTTACGTAGGGGACAGCAAGAGACCATGGAT 1253
QY      401 GlyTyrThrSerGluIlePheMetGlyLysAsnThrIleValLeuHisAsnThrCys 420
DB      1254 GAGTACACTTCAGAGATATTCATGAGGTGGAGAAAGACACCATTTGTTGGACAAACATGC 1313
QY      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB      1314 CAGGATTCCTCTTACGTGCTCTTATCTTGGACTTGGCTCTTCTGCTAGACTCAGC 1373
QY      441 ThrArgIleGlnPheLysAlaGluAsnGlnGlyLysPheHisSerPheHisProValAla 460
DB      1374 ACTAGAAATCGAGTTTAAAGCTGAATAATGAGGGAATAATTCATCTCAACCCAGTTGCT 1433
QY      461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB      1434 ACCATCTTCAGTATACCTCAACCAAGGCTCTCTGTTCCACCGGTACACCAATGTGTGAT 1493
QY      481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetCysArgAlaCysValGlyLeuAla 500
DB      1494 GCATTGTCAAAGCAGCGTCAATGCTGAAAAACATTAAGAGGCTTGTGTTGATTGGCC 1553
QY      501 ProGluAsnAsnMetIleLeuGluTyrLys 510
DB      1554 CCAGAGAAATAATGATATCTCGAGTACAAG 1583
  
```

RESULT 3

BD075269 1533 bp DNA linear PAT 27-AUG-2002
LOCUS BD075269
DEFINITION Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid.
ACCESSION BD075269
VERSION BD075269.1 GI:22620872
KEYWORDS JP 2001519665-A/4.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1533)
AUTHORS Hitz,W.D. and Sebastian,S.A.
TITLE Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid
JOURNAL Patent: JP 2001519665-A 4 23-OCT-2001;
EI DU PONT DE NEMOURS AND CO
COMMENT OS Soybean line LR3
PN JP 2001519665-A/4
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ,SCOTT ANTHONY SEBASTIAN
PC C12N15/52,C12N15/82,C12N15/11,C12N9/90,A01H5/00 CC
Strandedness: Double;
Topology: Linear;
CC Soybean plant producing seeds with reduced
levels of raffinose
CC and phytic acid
FH Key Location/Qualifiers
FT CDS 1..1533.
Location/Qualifiers
1..1533
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:

Pred. No.:	6,31e-199	Length:	1533
Score:	2588.00	Matches:	501
Percent Similarity:	99.02%	Conservative:	4
Best Local Similarity:	98.24%	Mismatches:	5
Query Match:	98.37%	Indels:	0
DB:	6	Gaps:	0

US-10-718-952-12 (1-510) x BD075269 (1-1533)

QY 1 MetPheilegluaasnphelysvalgiuserproasnvallyserthrighlurthrgluile 20
DB 1 ATGTTTCATGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
QY 21 GlnserValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCCGCTGTACAACTACGAAACCAACCGAACTGTTCACAGAAACGAGATGCACTAT 120
QY 41 GlnTPrileVallyserProlyserValAsnTyrGlnPheLySerThrAsnThrHisValPro 60
DB 121 CAGTGATTTGCAAAACCAAACTGTCAATACGAATTTAAACCAACATCCATGTTCTT 180
QY 61 LysLeuGlyValMetLeuValGlyTPrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 181 AAATTAGGGGTATATGTTGTGGGTTGGGGTGGAAACACGCGTCAACCTCACCGGTGTT 240
QY 81 ValIleAlaAsnArgGluAspIleSerTPrAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 241 GTTATTTGCTAACCGAGAGGCAATTTTCATGGGCTACAAAGACAAAGATTCACAAAGCCAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlnGly 120
DB 301 TACTTTGGCTCCCTCACCAAGCTCAGCTATCCGAGTTGGGCTCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140

DB 361 ATCTATGCCCATTCAGAGAGCTGCTTCACATGTTTAACTCTTGACGAGATGTGTTGGG 420
QY 141 GlyTPrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGGATATCAGCAACATCAACCTGCTATGCTCAGGCGCAAGAGTGTATAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGlySerMetValProLeuProGlyIle 180
DB 481 ATCGATTTGCAAGACAGATGAGGCTTACATGGAATCATGCTTCCACTCCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluValAlaAsnAsnValIleLysGly 200
DB 541 TATGACCCGAGTTTCATTCATTCGCAACCAAGAGAGCTGTCAACAGTATCAAGGCT 600
QY 201 ThrLysGlnGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
DB 601 ACNAGCAAGAGCANGTTCAACAAATCATCAAGACATCAAGGCTTTAAGAAACCAAC 660
QY 221 LysValAspLysValValIleuTPrThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGCAAGAGTGTGTGTACTGTGACTGCCAACACAGAGAGGTACAGTAATTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaAlaValAspArgAsnGluAlaIu 260
DB 721 GTGGGCTTAAATGACACCATGAGAAATCTTGGCTGTGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTCCTTCACCTTTATGATGCAATTCCTGTTGTATGAAATGTTCTTCTTCAATAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAGAACACTTTTGTACAGAGCTGATGATGATCTTGCCATGCGCAGGAAC 900
QY 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
DB 901 TTGATTGGTGAGATGACTTCAGAGTGTGACAGCAAAATGAAATCTGTGTGGTTAT 960
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 961 TTCCTTGGGGGCTGGATGATCAAGCAACATCTATAGCTATACCACTGGAACATCTGG 1020
QY 341 AsnAspGlyMetCysLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
DB 1021 AATGATGATATGAATCTTTCGCTCCCAAACTTCCGTTCAAGGAATCTCCAAAGAGC 1080
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlnHisPro 380
DB 1081 AACGTTGTGATGATATGATGTCACACGAAATGCCATCTATGAGACCTGTGTGAACAT 1140
QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCATGTTGTTATTAAGTATGTCCTTACGTAGGGGACAGAAATAGACCTGAT 1200
QY 401 GlnTyrThrSerGlnIlePheMetGlyGlyLysAsnThrIleValIleAsnThrCys 420
DB 1201 GAGTACACTTCAGAGATATTCATGCGTGAAGAAGCACCATTTGTTTCCAAACACATGC 1260
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTCTTTCGCTGCTCTTATTAATCTTGGACTTGTCCTTTCGAGCTGAGC 1320
QY 441 ThrArgIleGlnPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTAGAAATGATTAAGCTGAATAATAGGAAATATCCATCTATCCACCCCACTTCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB 1381 ACCATCTCAGCTACCTACCAAGGCTCTCTGTGTCCACCGGATGACCAAGTGTGAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGlnAsnIleMetArgAlaCysValGlyLeuAla 500

DB 1441 GCATTGTCAAGACGAGCGTCGAATGCTGAAGAAACATATAGAGGGCTTGTTGGATTGGCC 1500

QY 501 ProGUaenAaSmMeTleuGluTyrIle 510
|||||
DB 1501 CCAGAGAAATACATGATCTCGAGTACCAAG 1530

RESULT 4
AF293970
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

AF293970 1791 bp mRNA linear PLN 03-MAY-2001
Glycine max myo-inositol-3-phosphate synthase (MIPS) mRNA, complete cds.
AF293970 GI:13936690

Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 1791)
Hegeman,C.E., Good,L.L. and Grabau,E.A.
Expression of D-myo-Inositol-3-phosphate synthase in soybean.
Implications for phytic acid biosynthesis
Plant Physiol. 125 (4), 1941-1948 (2001)
21196082
11299373
2 (bases 1 to 1791)
Hegeman,C.E., Good,L.L. and Grabau,E.A.
Direct Submission
Submitted (07-AUG-2000) Plant Pathology, Physiology, and Weed Science, Virginia Tech, Fraulin Biotechnology Center, Blacksburg, VA 24060, USA

FEATURES
source
Location/Qualifiers
1. 1791
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
/dev_stage="Immature cotyledons"
1. 1791
/gene="MIPS"
62. 1594
/gene="MIPS"
/note="isomerase"
/codon_start=1
/product="myo-inositol-3-phosphate synthase"
/protein_id="AAK49896.1"
/db_xref="GI:13936691"
/translation="MPTMRKVCPCNVKKTETETIOSVNVYETTELVDHNRNCTQWIV
KRKSVKYEKTRNIHVPKGVMVGVGNNGSTLTGTVIANREGISWATKDKIQANVF
GSLTASAIKRVGSFGSEETVAPFKSLIPVNPDDIVFGMDISNNLADAMARAFVD
IDLQKMPYMSMLPLPGIYDPDFIAAOBERANNVIGMTQEOVOQIIXDKIKAKE
ATKRVKVVLTAMTERYSNLVVGNDIVENLLAANDREASISSTLYALICAMENV
PINSPOINTFVPGIDLAIAIRNTLIGDDPFKSGOTKMSVAVDLVAGAIKPTISVS
YNHLCNNNGMNTSABOTPRSKISKSNDVMSNALLIYEGCHPDHVVYIKYIPYV
GDSKRAMBYTSSEITMGKSTTVLHNTCCDSLAAPIIDLVLLAEVLSRIFSKANE
KPFHSFHPVATLSTYTRAPVLPSTPGVNALSKQRAMLENTMRAQCVGLAPENNMLE
YK"

ORIGIN

Alignment Scores:
Pred. No.: 3.37e-198 Length: 1791
Score: 2580.00 Matches: 500
Percent Similarity: 98.82% Conservative: 4
Best Local Similarity: 98.04% Mismatches: 6
Query Match: 98.06% Indels: 0
DB: 8 Gaps: 0

US-10-718-952-12 (1-510) x AF293970 (1-1791)

QY 1 MetPheIleGUaenPheIyValGluSerProAenValIySerThrGluThrgluIle 20
|||||
|||||

DB 62 ATGTCATCGAGAAATTTTAAGGTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATT 121

QY 21 GlnSerValTyrAenTyrGluThrThrgluLeuValHisGluAaenArgaenGlyThrTyr 40
|||||
DB 122 CAGTCGCTTACAACTACCAACCAACCGAACTGTTCACAGAAACGGAATGCGACCTAT 181

QY 41 GlnTPRILEValIySerProIySerValAenTyrGlnPheIySerThraenThrHisValPro 60
|||||
DB 182 CAGTCGATTTGCAACCAACCAATCTGTCAATACGAATTTAAACCAACATTCACATGTTCTT 241

QY 61 IySLeuGlyValIyMeLeuValGlyTyrGlyGlyAaenGlySerThrLeuThrglyGly 80
242 AATTAGGGGTATATGTTGTGGGTGGGATGGAACCAACGGCTCAACCTCCACGGTGGT 301

DB 81 ValIleAlaAaenArgGluAaenIleSerTPRAlaThrIyAaenIyIleGlnIleAlaAaen 100
|||||
DB 302 GTTATATGCTTACACCGAGGGCAATTCATGCGCTTACCAAGCAAGATTCACAAAGCAAT 361

QY 101 TyrPheGlySerLeuThrglnIleAaenAlaIleArgValGlySerPheGlnGlyGlu 120
362 TACTTTGGCTCCCTTACCCAGCCTCAGCTTATCCAGTTGGGTCTTCCAGGAGAGAA 421

DB 121 IleTyrAlaProPheIySerLeuLeuProMetValAaenProAaenIleValPheGly 140
422 ATCTAATGCCCATTCAGAGCCTGCTTCCAAATGTTAACCTCGACGACATTTGTTGGG 481

QY 141 GlyTPRAspIleSerAaenMetAaenLeuAlaAaenAlaMetAlaArgAlaIyValPheAaP 160
482 GGATGGAGATATCAGCAACATGAACTGCTGATGCTATGCCATGCCAGGCAAAAGGTGTTGAC 541

DB 161 IleAaenLeuGlnIyGlnLeuAaenProTyrMetGluSerMetValProleuProGlyIle 180
542 ATCGATTTCAGAAAGAGAGGAGGCTTACATGAAATTCATCTTCCTCCCGGAATC 601

QY 181 TyrAspProAaenPheIleAlaIleAaenGlnGluArgAlaAaenValIleIySgly 200
602 TATGACCCGGATTTCATTTGCTGCCAACCAAGAGGCGTCCCAACACTCATCAAGGCG 661

DB 201 ThrIySglnGlnIyValGlnGlnIleIleIyAaenIleIyAlaPheIySgluAlaThr 220
662 ACMAAGCAAGCAAGTTCACAAATCATCAAGATCATCAAGCGCTTAAAGCAAGCCACC 721

QY 221 IySValAaenIyValIyValIleuThrTPRAlaAaenThrgluArgIySerAaenIyVal 240
722 AAGTGGACAAAGTGTGTACTGTGACCTGCCAACACAGAGGATCAATTTGGTT 781

DB 241 ValGlyLeuAaenAspThrMetGluAaenLeuAlaIleValAaenIyAaenIyAaenIyAaen 260
782 GTGGGCTTAAAGACACCATGAGAAATCTTGGCTGCTGTGACAGAAATGAGCTGAG 841

QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAaenValProPheIleAaen 280
842 ATTTCTCTTCCACCTTGATCCATCTGCTGTGTATGGAAGAAATGTTCTTCAATTAAT 901

DB 281 GlysSerProGlnAaenThrPheValProGlyLeuIleAaenIleAlaIyAaenThr 300
902 GGAACCCCTCAGAAACATTTGTATCCAGGCGGATGTGATTCGCAATCCCGAGAAACAT 961

QY 301 LeuIleGlyGlyAaenAspPheIySerGlyInThrIyMetIySerValIleuValAaP 320
962 TTGATTGGTGAAGATTTTCAAGAGTGTCTGACCAAAATGAATTCGTGTGGTTGAT 1021

DB 321 PheLeuValGlyAlaGlyIleIySProThrSerIleValSerTyrAaenHisleuGlyAaen 340
1022 TTCCTTGGGGGCTGATATCAAGCAACATCTATAGTCAGTTAACACCATCTGGGAAC 1081

QY 341 AaenAspIyMetAaenLeuSerAlaProGlnThrPheAaenSerIyGluIleSerIySer 360
1082 AATGATGGATGAATCTTTCGGCTCCAAACTTTCCTTCAAGAAATCTCCAAAGAC 1141

DB 361 AaenValIleAaenAspMetValAaenSerAaenAlaIleuTyrGluProGlyGluHisPro 380
1142 AACGTTGTGATGATGTGTCACAGCAATGCAATCCTCTATGACCTGTGTGAATCA 1201

QY 381 Asph1ValValValIleuSerTyValProTyValGlyAspSerLysArgAlaMetAsp 400
DB 1202 GACCATGTTGTTGTTATTAAGATATGCTTTCGTAAGGAGACAGCAAGACCATGAT 1261
QY 401 GlnTyThrSerGluIlePheMetGlyGlyLysAsnThrIleValIleuHisAsnThrCys 420
DB 1262 GAGTACACTTCAGATATATTCATGCGTGGAAAGACACCATGTTGTCACAACACATGC 1321
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1322 GAGGATTCCTCTTGAAGCTGCTCTATATCTTGAGCTTGCTCTTCTTGCTAGCTCAGC 1381
QY 441 ThrArgIleGlnPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
DB 1382 ACTAGATCGACTTTAAGCTGAATGAGGAAATTCATCATCTTCACACCATGCTGT 1441
QY 461 ThrIleLeuSerTyLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
DB 1442 ACCATATCTCAGCTACCTCACCAAGCTCCTCTGTTCCACCGGTAACACCGGTTGAT 1501
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1502 GCATTGTCAGACGCGTCAATGCTGGAACATTAATGAGGCTTGTTGTTGATTGGCC 1561
QY 501 ProGluAsnAsnMetIleLeuGluTyTyLys 510
DB 1562 CCAGAGATTAACATGATTTCTCGAGTACAAAG 1591
RESULT 5
LOCUS AB009881 1959 bp mRNA linear PLN 13-JUL-2000
DEFINITION Nicotiana tabacum mRNA for myo-inositol 1-phosphate synthase,
complete cds.
ACCESSION AB009881
VERSION AB009881.1 GI:8096265
KEYWORDS myo-inositol 1-phosphate synthase.
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; Lamiales; Solanales; Solanaceae; Nicotiana.
REFERENCE
AUTHORS Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.
TITLE Screening of wound-responsive genes identifies an immediate-early
expressed gene encoding a highly charged protein in mechanically
wounded tobacco plants
JOURNAL Plant Cell Physiol. 41 (6), 684-691 (2000)
MEDLINE 20399434
PUBMED 10945337
REFERENCE
AUTHORS Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.
TITLE Direct Submision
JOURNAL Submitted (18-DEC-1997) Kojiro Hara, Nara Institute of Science and
Technology, Plant Molecular Breeding, Takayama 8916-5, Ikoma, Nara
630-0101, Japan (E-mail:k-hara@bs.aisf-nara.ac.jp,
Tel:+81-743-72-5653 (ex.5653), Fax:+81-743-72-5659)
FEATURES
source
1..1959
/organism="Nicotiana tabacum"
/mol_type="mRNA"
/db_xref="taxon:4097"
/tissue_type="leaf"
94..1626
/product="myo-inositol 1-phosphate synthase"
/protein_id="BA095788.1"
/db_xref="GI:8096266"
/translation="MPIENFKVESPVNVKTESIHSVYDYOTTELVEHEKNGTYOMTV
KPKTVYERKDVHVPKGLVWLVGWNGSGSLTGGLVIANREGISMATKDKYQANVF
GSULTASTIRVSGFNGEBIYAFPKSLILPVNPDDVYFGSMDISGNLADANARAFD
IDLQKLRPKYSWVPLPEITYPDPLAAGSRANNAVKGITKEEDIDDIITKIRFEKE
KNKDKVVLWMTANTERSNVVVGINDTWNELFASVDRNEAIEISPTIYALACILENV

polyA_site
ORIGIN
1959
/note="15 a nucleotides"
Alignment Scores:
Pred. No.: 1,076-189 Length: 1959
Score: 2475.00 Matches: 473
Percent Similarity: 96.27% Conservative: 18
Best Local Similarity: 92.75% Mismatches: 19
Query Match: 94.07% Indels: 0
Gaps: 0
US-10-718-952-12 (1-510) x AB009881 (1-1959)
QY 1 MetPheIleGluAsnPheLysValGluSerProAsnValLysTyThrGluThrGluIle 20
DB 94 ATGTTATTTGAAGACTTTAAGTTGAGAGCCCAACGTTAAGTACACCGAAAGTGAGATT 153
QY 21 GlnSerValIlyrAsnTyGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTy 40
DB 154 CACTCTGCTATGATTTATCAACCACTAGTTAGTTCTATGAGAGAGAAAATGGGACTTAC 213
QY 41 GlnTrpIleValLysProLysSerValAsnTyGlnPheLysThrAsnThrHisValPro 60
DB 214 CATGAGCTGTCACACCTTAAGACTCTCAAAATATGATGTTCAAGACTGATCTTATTTCC 273
QY 61 LysLeuGlyValMetLeuValGlyTyProLysGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 274 AATTAGGGGTTATATCTTGTGATGGGTGGAACATATGTTCAACTTGAACCGTGGT 333
QY 81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
DB 334 GTTATTTGCTAACAGAGAAAGAAATTCATAGGCGCAACAAAGATTAAGTCCACACCAAT 393
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGly 120
DB 394 TACCTTGCTCTTCTTACTCAGGCTTCTACTTTCAGATTTGGGCTTTCAATGGAGAARG 453
QY 121 IleTyAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 454 ATCTATGCTTCATTTAAAGCTCCTCTTCAATGTCATCCAGATGAGTGTGTTGGA 513
QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 514 GGATGGACATCAGCGGCATGATTTGGCAGATGCGCAGAGGCTTAAGGATTTGAT 573
QY 161 IleAspLeuGlnLysGlnLeuArgProTyMetGluSerMetValProLeuProGlyIle 180
DB 574 ATTGATCTCAAAAGCAAGTTAGAGCCCTTACATGAAATCTATGCTCCACTACCTGATTC 633
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 634 TATGATCGTGAATTCATTTGCTGCTTACCAAGGATCAGCTGCCAACAGTGTATCAAGGA 693
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 694 ACCAAGAAAGAAACAAATGATCAATTCATTAAGATATTAGGAGATTTAAGAGAAAGAAC 753
QY 221 LysValAspLysValValIleuThrThrAlaAsnThrGluArgTySerAsnLeuVal 240
DB 754 AAAGTGAACAGAGTGTGCTTGTGACCTGTTACACTGAAGATACAGTATATGGTT 813
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsnGluAlaGlu 260
DB 814 GTTGACCTTAATGATACATATGAAATCTCTTGTCTGTGGAACAGAAATCAACTGAA 873
QY 261 IleSerProSerThrLeuTyAlaIleAlaCysValMetGluAsnValProPheIleAsn 280

(E-mail: Akiko.Hashimoto@pbgrl.jti.co.jp, Tel: +81-538-32-7116,
Fax: +81-538-33-6046)

DB 874 ATATCTCTCCACTTTGTATGCTATTGGCTGCAATCTTGAAGAATGCTTTATTAT 933

QY 281 GYSerProGlnAanthrPheValProGlyLeuIleAspLeuAlaIleAlaArgAnthr 300

DB 934 GGAAGCCCAAGAACCTTTGTGCCAGGCTCATTTGATTTGGCCATTCAGAGAACAA 993

QY 301 LeuIleGlyAspAspPheLeuSerGlyGlnThrLeuMetLeuSerValLeuValAsp 320

DB 999 CTGATTGGTGGTGAATCTTTAGAGTGGTCAACCAATGAGTCACTGCTGTTGAT 1053

QY 321 PheLeuValGlyAlaGlyIleLeuProThrSerIleValSerTyrAsnHisLeuGlyAsn 340

DB 1054 TTCCTTTGTTGAGCTGTATTAAGCAACATCAATTTGAGGTACAAACATTTGGTAA 1113

QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLeuGlyIleSerLeuSer 360

DB 1114 AATGATGGAATGATCTGCTGCTCCCTCAAACTTCCCTCAAGAGAGATCTGAAAGT 1173

QY 361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyGlnHisPro 380

DB 1174 AATGTTGTTGATGACATGATTTCAAGCAATGCCATTTCTTAAGGCTTGAAGACCT 1233

QY 381 AspHisValValIleLeuSerTyrValProTyrValGlyAspSerLeuArgAlaMetAsp 400

DB 1234 GACCATGTTGTTGTATTAAGTATGTCATATGTCGAGACAGCAAAAGGCAATGAT 1293

QY 401 GlnTyrThrSerGlnIlePheMetGlyIleValAsnThrIleValLeuHisAsnThrCys 420

DB 1294 GAGTACACATCTGAGATTTTCATGAGAGAGAAAGAACACATTTGTTGCAACATCTGT 1353

QY 421 GlnAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGlnLeuSer 440

DB 1354 GAGGATCTCTTTGGCTCTCCAAATTATCTTGATTTGGTCTTTCTTCTGAACTCAAT 1413

QY 441 ThrArgIleGlnPheLeuValGlnAsnGlnGlyIlePheHisSerPheHisProValAla 460

DB 1414 ACCCCCATTCACGCTCAAACTGAGAGAGAGGTAAGTTCCATCTCTTCCACCCCGGCT 1473

QY 461 ThrIleLeuSerTyrLeuThrIleValAlaProLeuValProProGlyThrProValValAsn 480

DB 1474 ACTATCTTCACGCTTACCTTCCAGAGCTCTCTGTCATCCACAGGTACACAGGTGGTAA 1533

QY 481 AlaLeuSerLeuGlnArgAlaMetLeuGlnAsnIleMetArgAlaCysValGlyLeuAla 500

DB 1534 GCACCTTCAGAGCAAGGCGCAATGCTTGAGAACATATGAGGCTTGTGTGACTTGA 1593

QY 501 ProGlnAsnAsnMetLeuGlnTyrIle 510

DB 1594 CCAGAGAACACATGATTTCTGGAATACAA 1623

RESULT 6

AB032073 1931 bp mRNA linear PLN 08-SEP-1999

LOCUS Nicotiana paniculata NpINP51 mRNA for myo-inositol-1-phosphate

DEFINITION synthase, complete cds.

ACCESSION AB032073.1 GI:5834499

VERSION myo-inositol-1-phosphate synthase.

KEYWORDS Nicotiana paniculata

SOURCE Nicotiana paniculata

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE Hashimoto, A., Yamada, S. and Komori, T.

AUTHORS Myo-inositol-1-phosphate synthase

JOURNAL Published Only in Database (1999)

REFERENCE 2 (bases 1 to 1931)

AUTHORS Hashimoto, A., Yamada, S. and Komori, T.

TITLE Direct Submission

JOURNAL Submitted (01-SEP-1999) Akiko Hashimoto, Japan Tobacco Inc., Plant

Breeding and Genetics Research Laboratory; 700 Higashibara, Iwata,

Toyoda-cho, Shizuoka 418-0802, Japan

FEATURES

source Location/Qualifiers

1..1931

/organism="Nicotiana paniculata"

/mol_type="mRNA"

/db_xref="taxon:62141"

1..1931

/gene="NpINP51"

/gene="NpINP51"

92..1624

/gene="NpINP51"

/product="myo-inositol-1-phosphate synthase"

/codon_start=1

/protein_id="BA84084.1"

/db_xref="GI:5834500"

/translation="MFLENFKESPNVRYTSEHSIVDYQTTLEVDEKNGTYQMTV
KPKTYKFEKIDVHPKGLVIMVGWGNNGSLTGGVLANRGIEMATDQKQANYP
GSLTQASTLRVGSFNGSEIYAPFKSLPMVNPDDVYFGMDISDMNLADMAARAFD
IDLOKQLPYESWVPLPGIYDDPFIANQGSRAANVYKGTKEQIDQIKIDREKE
KKKDVVVVLTMTANTERSNVYVGNPTMENIPASVDNREISSTLYAICILENV
PFINSPQNTFVPGILDAIKKNTLIGDDFRSGQTKSVLVDLVAGIKPTISV
YNHLDGNDGMNLAPQTRSKSISKNVDDWVSSNALYERGERHDVVVYIKYVYV
GDSKRAMDEYSEIFMGKNTIIVLHNTCEDSLAIPILDLVLAELESTRDLKAGE
GKFSHFHVATILSYLTKAPLVPGTPIVNALSKRAMLENILRACVGLAPENNMLE
YK"

ORIGIN

Alignment Scores:

Pred. No.: 2,656-189 Length: 1931

Score: 2470.00 Matches: 471

Percent Similarity: 96.47% Conservative: 21

Best Local Similarity: 97.35% Mismatches: 18

Query Match: 93.88% Indels: 0

DB: 8 Gaps: 0

US-10-718-952-12 (1-510) x AB032073 (1-1931)

QY 1 MetPheIleGlnAsnPheLeuValGlnSerProAsnValIleTyrThrGlnThrGlnIle 20

DB 92 ATGTTATTGAGAAATTTTAAGGTGAGAGCCCAAGTTAATGATACACCGAAGTAAT 151

QY 21 GlnSerValTyrAsnTyrGlnThrThrGlnLeuValHisGlnAsnArgAsnGlyThrTyr 40

DB 152 CACTCTGCTATGATTAACAACCTGAGTTAGTTCAATGAGAAAAATGGACATAT 211

QY 41 GlnTyrIleValIleAspProLeuSerValAsnTyrGlnPheLeuSerThrAsnThrHisValPro 60

DB 212 CAATGACCGCTCAAGCCTTAAGCTGCAAAATATGAGTTCAAGACTGATGTCATGTTCCC 271

QY 61 LysLeuGlyValMetLeuValGlyTyrGlyIleValAsnGlnGlySerThrLeuThrGlyGly 80

DB 272 AAATTAGGGGTATATCTTTGGATGGGTGGAAACAAATGTTCAACTGACCGGTGCT 331

QY 81 ValIleAlaAsnArgGlnAspIleSerTyrAlaThrIleAspValIleGlnGlnAlaAsn 100

DB 332 GTTATTGCTTACAGAGAAATTTCAAGGCCACCAAGATTAAGGTGCAACAGCAAT 391

QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlnIle 120

DB 392 TACTTTGGGCTCTTACTACAGGCTTCTACTATTCAGAGTTGGGCTTTCAATGAGAGAG 451

QY 121 IleTyrAlaProPheLeuSerLeuLeuProMetValAsnProAspAspIleValPheGly 140

DB 452 ATCTATGCTCCATTTTAAAGCTCTTCCAAAGTCAATCCAGATGAGAGCTACTGTTTGA 511

QY 141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValPheAsp 160

DB 512 GGATGGGACATCAGGACATGATTTAGCATGTCAGGCGCAAGGCTTAAGTATTTGAT 571

QY 161 IleAspLeuGlnValGlnLeuArgProTyrMetGlnSerMetValProLeuProGlyIle 180

DB 572 ATGATCTACAAAGCAAGTGGGCTTCAATGAAATCTATGTCCTCCACTCGTGTATC 631

QY 181 TyrAspProAspPheIleAlaIaIaengIngluArgAlaAsnAsnValIleYsgly 200
DB 632 TATGACCTGATTTATCTGCTCTAACCAGGCTCACGCGCAACAACGTGATCAAGA 691
QY 201 ThrIysGIngluGlnValGlnGlnIleIleYsAspIleYsAlaPheYsGluAlaThr 220
DB 692 ACCAAGAAAGAAACAATTGATCAATTAAGATTAATAGGAGCTTTAAGAAAAAGAC 751
QY 221 LysValaAspLysValaValaValleuThrPheIleAsnThrGluArgTyrSerLeuVal 240
DB 752 AAAGGAGCAAGGTGTGATGTGAGACTGTCAACAGTGAAGATACGATATGTGGTT 811
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsnGluAla 260
DB 812 GTTGACCTTAATGACATATGAGAAACCTTTGCTCTGTGACAGAAATGAACTGA 871
QY 261 IleserProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 872 ATATCTCCTTCCACTTGTATGCTATGTGCTGATCTTGAATAATGTGCTTTATTAAT 931
QY 281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 932 GGAAAGCCCCAGAACACCTTGTGCCAGGCTCATTTGAGCATCAAGAAACACA 991
QY 301 LeuIleGlyGlyAspAspPheYsSerGlyGlnThrIysMetYsSerValLeuValAsp 320
DB 992 TTGATTTGGTGGTGAACCTTTAAGAGTGTCTCAACAAATGAATCAAGTGTGTTGAT 1051
QY 321 PheLeuValGlyAlaGlyIleYsProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 1052 TTCCCTGTGTGAGCTGTATTAAGCCAACTCAATTTGAGCTACACCATTTGGGTAAAC 1111
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerGlnIleSerYsSer 360
DB 1112 AATGATGGATAAATGTCTCTCCCTCAACTTCCTCGGTCAAGAGATCTGAAAGAT 1171
QY 361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1172 AATGTTGTTGATGACATGCTTCAAGCAATGCCATCTTTAAGAGCTGAGAGACCTT 1231
QY 381 AspHisValAlaValIleYsTyrValProTyrValGlyAspSerIleYsAlaMetAsp 400
DB 1232 GACCAATGTTGTGTATTAAGTATGATGCGCATGTGTGAGACAGCAAGGCGCAATGAT 1291
QY 401 GlyTyrThrSerGlnIlePheMetGlyGlyYsAsnThrIleValLeuHisAsnThrCys 420
DB 1292 GAGTACACATCTGAGATTTTCAAGGGGAGAAACACCATTTGTTGCAATATCTGT 1351
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1352 GAGGATTCATCTTTAGCTGCTCCAAATTAATTTGATTTGCTCTTCTGTAACCTGAT 1411
QY 441 ThrArgIleGlnPheYsAlaGluAsnGluGlyYsPheHisSerPheHisProValAla 460
DB 1412 ACCCCATTCACATCAAGCTCAAGGAGGAGGATTAATTTGCTCTTCCACCCCGTGCT 1471
QY 461 ThrIleLeuSerTyrLeuThrIleYsAlaProLeuValProProGlyThrProValAsn 480
DB 1472 ACTATCTCAGCTACCTTAACCAAGGCTCTCTGTGACCAACAGGTACACCAATGTGGAT 1531
QY 481 AlaLeuSerYsGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1532 GCACCTCTCAAGCAAGGCGCAATGCTTGAGAACATATTAAGGGCTTTGTGTGACTTGA 1591
QY 501 ProGluAsnAsnMetIleLeuGlyTyrIys 510
DB 1592 CCAGAGAACCAACATGATCTGGAATACAA 1621
RESULT 7
LOCUS E27176 1950 bp DNA linear PAT 18-JUN-2001
DEFINITION Novel INPS gene derived from nicotiana.
ACCESSION E27176

VERSION E27176.1 GI:13026394
KEYWORDS UP 1999187879-A/1.
SOURCE Nicotiana paniculata
ORGANISM Nicotiana paniculata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; Lamiales; Solanales; Solanaceae; Nicotiana.
REFERENCE Shigehiro, Y. and Toshiyuki, K.
AUTHORS Novel INPS gene derived from nicotiana
TITLE Patent: UP 1999187879-A 1 13-JUL-1999;
JOURNAL JAPAN TOBACCO INC
COMMENT OS Nicotiana paniculata
PN UP 1999187879-A/1
PD 13-JUL-1999
PF 26-DEC-1997 UP 1997359773
PR SHIGEHIRO YAMADA, TOSHIYUKI KOMORI
PI C12N15/09//C12N5/10, (C12N15/09, C12N1.91), C12N15/00, C12N5/00,
PC (C12N15/00, C12N1.91)
CC Strandedness: Double;
CC Topology: Linear;
PH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..1950
/organism="Nicotiana paniculata"
/mol_type="genomic DNA"
/db_xref="taxon:62141"
ORIGIN
Alignment Scores:
Pred. No.: 2,696-189 Length: 1950
Score: 2470.00 Matches: 471
Percent Similarity: 96.47% Conservative: 21
Best Local Similarity: 92.35% Mismatches: 18
Query Match: 93.88% Indels: 0
DB: 6 Gaps: 0
US-10-718-952-12 (1-510) x E27176 (1-1950)
QY 1 MetPheIleGluAsnPheYsValGluSerProAsnValIleYsTyrThrGluThrGluIle 20
DB 92 ATGTTATTTAGAAATTTTAAGTTGAGAGCCCAACGTTAAGTACCCGAAGAATGAATTT 151
QY 21 GlnSerValIleYsAsnTyrGluThrThrGluLeuValIleGluAsnArgAsnGlyThrIyr 40
DB 152 CACTCTGCTATGATTAATCAACCACTGAGTTAGTTCAATGATGAGAAAAATGGGACATAT 211
QY 41 GlnTPrIleValIleYsProYsSerValAsnTyrGlnPheYsThrAsnThrHisValPro 60
DB 212 CAATGACCGCTCAAGCCCTTAAGACTCTCAAAATTAAGATTCAAGACTGATGATCTTCC 271
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 272 AAATTAGGGGTTATGCTTTGATGAGTGGGTGAACCAATGGTTCAACCTTGACCGGTGT 331
QY 81 ValIleAlaAsnArgGluAspIleSerTyrAlaThrIleYsAspLysIleGlnGlnAlaAsn 100
DB 332 GTTATTTGCTAACAGAGAAAGAAATTTCAATGGGCGCACCAAGATTAAGTCCAAAGCCAT 391
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValIleYsSerPheGlnGlyGluGlu 120
DB 392 TACTTTGGCTCTCTTACTCAAGCTTTCTACTATTCGATGTGGGTCTTTCAATGGAGAAAG 451
QY 121 IleTyrAlaProPheYsSerLeuLeuProMetValAsnProAspAspIleValaPheGly 140
DB 452 ATCTATGCTCATTTAAAGCCTCTTCCAAATGCTCAATGCAATGCAATGCAATGCTGTTGA 511
QY 141 GlyTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleYsValPheAsp 160
DB 512 GGATGGACATCAAGCAGATGAATTTTACAGATGCCATGGCCAGGGCTTAAGGTATTTGAT 571

```

OY 11eapleuglnlysglnleuargProGlyrMetGluSerMetValProleupProGlylle 180
DB 572 ATTGATCTCAAAAGAGATTGAGCCCTTACATGAAATCTATGATCCCACTCCCTGGTATC 631
OY 181 TTAAspProAspPheIleAlaAaenglngluuargAlaAsnAnValIlelysgly 200
DB 632 TATGACCTGATTTCTGCTGCTTACCAAGGCTCAGTGCACCAACAGTATCAAAAGA 691
OY 201 ThrlysglngluInValglnlnlleIlelyAspIlelyAlaPheIlysglyAlaThr 220
DB 692 ACCAAGAAAGAAACAAATGATCAATCAATTAAGAAATTAAGAGCTTTAAGAAAAAAC 751
OY 221 LysValAspLysValValleuTPThrAlaAsnThrGluArgThrSerAsnLeuVal 240
DB 752 AAAGTGGACAGGTGATGATTTGTGACTGCTTACACAGAAAGATACAGTATGTGTT 811
OY 241 ValGlyLeuAsnAspThrMetGluAsnLeuValAlaValAspArgAsnGluAlaGlu 260
DB 812 GTTGGACTTATGACCTATGAAAACTTTGCTTGTGTGACAGAAATGAGCTGA 871
OY 261 IleserProserThrleuTyralIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 872 ATATCTCCTTCCACTTGTATGCTATGCTGCTGATCTTGAAATGCTTTTATTAAT 931
OY 281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 932 GGAAGCCCCAGAACACCTTTGTCCAGGCTCATGATTTGGCAATCAAGAAAGAACACA 991
OY 301 LeuIleGlyGlyAspAspPheIlySerGlyGlnThrIleMetLysSerValleuValAsp 320
DB 992 TTGATTGGTGGTGATACCTTTAAGAGTGTCAAAACAAATGAAGTCAAGTGGTGTGAT 1051
OY 321 PheLeuValGlyAlaGlyIlelySProThrSerIleValSerTyraAsnHisleuGlyAsn 340
DB 1052 TTCTCTTGTGGAGCTGGTATTAAGCCAAATGATGACATCAATGATGACCAACATTTGGGTAAC 1111
OY 341 AAspAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIlysglnIleSerLysSer 360
DB 1112 AATGATGAAATGATATGCTGTGCTCCCTCAAACTTCGGTCAAAAGAGATCTCGAAAGAT 1171
OY 361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrgIuProGlyGlnHisPro 380
DB 1172 AATGTTGTTGATGACATGTTTCAACCAATGCCATCTTTATGAGCTGGAAGGACCTT 1231
OY 381 AspPheValValIlelyStryValProTyryValGlyAspSerIlyAsnArgAlaMetAsp 400
DB 1232 GACCAATGTTGTGTGATTAAGTATGTCATATGTGGAGACAGCAAGAGGGCAATGGAT 1291
OY 401 GlnTyrrThrSerGlnIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys 420
DB 1292 GAGTACACATCTGAGATTTTCATGGGGGAGAAAGAACACAAATGTTTTCACAAATCTTGT 1351
OY 421 GlnAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1352 GAGGATTCATCTTTGAGCTGCTCCAAATTAATTCGATTTTGTCTTTCGTAACCAAT 1411
OY 441 ThrArgIleGlnPheIlyAlaGluAsnGlyGlyLysPheHisSerPheHisProValAla 460
DB 1412 ACCCGCATTCACGCTCAAAAGCTGAAGAGAGGGTAAATTCACCTCTTCCACCCCGGAGCT 1471
OY 461 ThrIleLeuSerTyrrleuThrIlyAsnAlaProleuValProProGlyThrProValAlaAsn 480
DB 1472 ACTATCTCAGCTACCTTACCAAGGCTCTCTGTGATACCAACAGATACACAGTGGGAAAT 1531
OY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1532 GCACTCTCAAAGAGAGGGCAATGCTTGAAGAACATTTGAGGGCTTGTGTGCACTTGA 1591
OY 501 ProGluAsnAsnMetIleleuGlyTyrrLys 510
DB 1592 CCAAGAACAAACATGATCTCGAATACAAA 1621

```

```

BT013505
LOCUS      BT013505                1959 bp      mRNA      linear      PLN 11-MAY-2004
DEFINITION      Lycopersicon esculentum clone 132193F, mRNA sequence.
ACCESSION      BT013505
VERSION        BT013505.1  GI:47104920
KEYWORDS
SOURCE
ORGANISM
  Lycopersicon esculentum (tomato)
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
  1 (bases 1 to 1959)
REFERENCE
  Kirkness,E.F., Mang,W. and Vazelle,A.
  Direct Submission
  Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
  Medical Center Drive, Rockville, MD 20850, USA
FEATURES
  source
    1..1959
    /organism="Lycopersicon esculentum"
    /mol_type="mRNA"
    /db_xref="taxon:4081"
    /clone="132193F"
    /cissue_type="fruit"
    /note="TZXBFS1"
ORIGIN
Alignment Scores:
Pred. No.:      3,92e-189      Length:      1959
Score:          2468.00      Matches:      472
Percent Similarity: 96.27%      Conservative: 19
Best Local Similarity: 92.55%      Mismatches:  19
Query Match:    92.80%      Gaps:        0
DB:
US-10-718-952-12 (1-510) x BT013505 (1-1959)
OY 1 MetPheIleGluAsnPheIlyValGluSerProAsnValIlySerThrGluThrGluIle 20
DB 77 ATGTTATATTAAGAAATTTTAAGGTGGAAGCCCAATATGAACTATATTCAGAAATGAATTT 136
OY 21 GlnSerValIlyrAsnTyrgIuThrThrGluLeuValHisGluAsnArgAsnGlyThrTyrr 40
DB 137 CATTCGTGTATGATTAAGAAACCAAGAGCTGTTCATGAAGAGAAATGAACTTAT 196
OY 41 GlnTPDlleValIlyeProLysSerValAsnTyrgIlnPheIlySerThrAsnThrHisValPro 60
DB 197 CAATGCAATTTTAACCTTAAACTGTCAAAATATGAATTTAAACTGATACCATGAGCCA 256
OY 61 LysLeuGlyValMetLeuValGlyTyrrGlyAsnAsnGlySerThrleuThrGlyGly 80
DB 257 AAATTGGGGGTATATCTTTGGATGAGGAGGAAACAATGTTCAACATTCAGTGAAGGT 316
OY 81 ValIleAlaAsnArgIluAspIleSerTPAlaThrIlyAspLysIleGlnGlnAlaAsn 100
DB 317 GTTATTCGCAATCGAAGAGATTTTCATGCGCAAGAAAGAAAGAAAGTGCACAAAGCCAT 376
OY 101 TyrrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGln 120
DB 377 TATTTTGGGTCTTAACTCAGGCAATCAACCATTCAGATTGGGTCTTTCATATGGCAAGAG 436
OY 121 IleTyrrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 437 ATCTATGACCCCTTAAAAAGCTCTTCCCATGTCACCCAGACGATGATATTTTGA 496
OY 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIlySerValPheAsp 160
DB 497 GGATGGACATTAAGCAATCATGATTTTGCAGATGCTATGCTAAGGGCTTAAGTTTTCGA 556
OY 161 IleAspLeuGlnIlyGlnLeuArgProTyrrMetGluSerMetValProleupProGlylle 180
DB 557 GTTGAATCGCAAAACAGCTGAGGCCCTTACATGAAATCATATGATCCCTCTCTGTATC 616
OY 181 TyrrAspProAspPheIleAlaAlaAaenglngluuargAlaAsnAnValIlelysgly 200

```

```

Db      617 TATGACCCGACTTCATTCGGCTAACCAAGACGCGCAACACCTGATCAAGA 676
Qy      201 ThrVSGInGluValGInGInIleIleYAspIleValAlaPheYSGInuIaThr 220
Db      677 ACCAAGAAAGAACCAAGTTCACCAAAATTTGTTAAAGATTTATTAAGGAGTTCAAGAGAAAGAC 736
Qy      221 LysValAspLysValValIleuTrpThrIaAsnThrGluArgTyrSerAsnLeuVal 240
Db      737 AAGGTAGACAGATAGTGGTTCTATGACTGCCCAACCGAAAGGTACAGTAAATGGT 796
Qy      241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluValGlu 260
Db      797 GTTGGCTTAATGACACCATGAAAACTTTTAGCTGCTGGATGAAATAGAGCTGAA 856
Qy      261 IleserProSerThrIleuTrpThrIaIleValCysValMetGluAsnValProPheIleAsn 280
Db      857 AATATCTCTTCTTACATTCATGCTATTCCTGCTTATATGAAAAAGTGCCTTTCATCAAC 916
Qy      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      917 GGAAGCCCTCAAAACACTTTTGTCCAGGCTTATGATTTGGCCATAAAGAAACACT 976
Qy      301 LeuIleGlyValAspAspPheYSGInuIleuTrpLysMetLysSerValIleuValAsp 320
Db      977 TTAATGTTGGTGTGATGACTTAAAGTGTCAACCAACAGATGAAGTCCGTGTGGTGTAT 1036
Qy      321 PheLeuValGlyValGlyIleLysProThrSerIleValSerTyrAsnHisIleuGluValAsn 340
Db      1037 TTCTCTTTGGAGCCGATTTAAAGCCACGTCATATGTAGCTCAATCAATCACTTGGGTATAC 1096
Qy      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db      1097 AATGATGAGATGATCTTCTGCTCCTCAACCTTCGCTTAAAGAGATCTCAAAAGAT 1156
Qy      361 AsnValValAspAspMetValAsnSerAsnAlaIleuTrpGluProGlyGluHisPro 380
Db      1157 AATGTTGTGATGACATGCTGTCTAGCAACGCCATCTCTATGAACTTGGCAGACACCT 1216
Qy      381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db      1217 GACCATGTTGTGTATCAAGATGTTCATATGTTGGAGACAGCAAGGGGCAATGAT 1276
Qy      401 GlnTyrThrSerGluIlePheMetGlyGlyLysAsnThrIleValIleuHisAsnThrCys 420
Db      1277 GAGTACATGTCAGAGATTTTTCATGGCGGAAAGACACTATAGTATGCACAAACCTTGT 1336
Qy      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleuLeuAlaGluLeuSer 440
Db      1337 GAGGACTCTCTTTGGACAGCTCCAAATTAATCTTGGAATTTGTCCTCTCCGTAACCTCACG 1396
Qy      441 ThrArgIleGlnPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db      1397 ACTCGCATTCAGCTCAAAAGCTGAAGGGAGGGGAGGAACTTCATCTCTTCCACCTCTGGGCT 1456
Qy      461 ThrIleLeuSerTyrIleuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db      1457 ACTATCTCTAGCTACTCAACCAAGGCTCCCTGGTATCCACCAAGTATCTCCAGTGGGAT 1516
Qy      481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db      1517 GCCCTTTCAAGAGAGGGCAATGCTTGAGAAATATATAGAGAGCTTGTGTGGTTGGCA 1576
Qy      501 ProGluAsnAsnMetIleLeuGluTyrLys 510
Db      1577 CCAGAGAACACATGATTTTGGAAATCAAA 1606

```

RESULT 9
 AF284065
 LOCUS
 DEFINITION Sesamum indicum myo-inositol 1-phosphate synthase mRNA, complete
 CDS.
 AF284065
 ACCESSION

```

VERSION AF284065.1 GI:3958815
KEYWORDS Sesamum indicum (sesame)
SOURCE Sesamum indicum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterids; Lamiales; Pedaliciales; Sesamum.
REFERENCE 1 (bases 1 to 1845)
AUTHORS Chun,J.A., Jin,U.H., Lee,J.W., Yi,Y.B., Hyung,N.I., Kang,M.H.,
Pyee,J.H., Suh,M.C., Kang,C.W., Seo,H.Y., Lee,S.W., and Chung,C.H.
Isolation and characterization of a myo-inositol 1-phosphate
synthase cDNA from developing sesame (Sesamum indicum L.) seeds:
functional and differential expression, and salt-induced
transcription during germination
Plant 216 (5), 874-880 (2003)
JOURNAL MEDLINE
PUBMED 22511750
REFERENCE 12624775
AUTHORS Jin,U.-H. and Chung,C.-H.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2000) Division of Biotechnology, Faculty of Life
Science and Resources, Dong-A University, Ha-Dan-Dong, Sa-Gu-Gu.,
Pusan City, Pusan 604-714, South Korea
FEATURES
source
1..1845
Location/Qualifiers
1..1845
/organism="Sesamum indicum"
/mol_type="mRNA"
/db_xref="taxon:4182"
/tissue_type="seeds"
/note="authority: Sesamum indicum L."
51..1583
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AA01148.1"
/db_xref="GI:3958815"
/translation="MPFESFVSESFNKYTEGEIHSVNYVETTELAVHESRNGTYOMIV
KEKTVKEFKTDHVPKLGKLVKMGKNGSTLTGCIYARREGISNAIKDYQANVF
GSLTQASSIRVGSNGEETIAPFKSLIPVAPDVVVGWDISNMNLADAGRAVLVD
IDLQKLRPVEHNVPLPGIYDPDFIANQSSRANNVIGTKKQVQDIIDKMDPFE
ONKDYKVVLTANTERSNVVGLINPAESILMASVERNEAIEPSTLYALACVEFN
PINGSPOINTVPGILIDAIORNSLIGDDFESQOTKXSVLVDLVGAGIKPTISVS
YNHGNDNGMLSPQPFRESKETSNSVVDVMVASNGILYRGRPHIVIKVPPVY
GDSKRADEYTSSEIFMGSKSTIVLANTCEBSLAPPIIDVLAEISTRQLARBE
GKFSFHEVATILSYLTAPLVPPGTPVVALSKORAMLENILRACVGLAENNMILR
YK"
ORIGIN
Alignment Scores:
Pred. No.: 7,06e-188 Length: 1845
Score: 2452.00 Matches: 470
Percent Similarity: 96.08% Conservative: 20
Best Local Similarity: 92.16% Mismatches: 20
Query Match: 93.20% Indels: 0
DB: 8 Gaps: 0
US-10-718-952-12 (1-510) x AF284065 (1-1845)
Qy 1 MetPheIleGluAsnPheLysValGluSerProAsnValLysTyrThrGluThrGluIle 20
Db 51 ATGTTCAATCGAGAGCTTCAAGGTTGAGAGCCCAATGTCAAGTACACCGAAGGTAGAT 110
Qy 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 111 CACTGTGCTCAACTACGAGACTTGTGTTCTATGAGAGAGAAAGGCACTTAC 170
Qy 41 GlnTrpIleValLysProLysSerValAsnTyrGlnPheLysTyrAsnThrHisValPro 60
Db 171 CAATGATGTTCAGCCCAAGACTGTCAAAATCGAGTTCAAAACCGATACATCATGCTCCC 230
Qy 61 LysLeuGlyValMetLeuValGlyTTrpGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db 231 AATTTAGGGGTATGCTTGTGTGGATGGGAGGAAACAATGGTTCAACTCTACCTGGCGGT 290

```

QY 81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnIleAlaAsn 100
 DB 291 GTTCATTGCCAAACAGGAGAGGATTTCATGGGCAACAAAGATTAAGTGCACAGGCCAAT 350
 QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValIleYserPheGlnIleGlu 120
 DB 351 TATTTTGGGTCTACTGACACAGGATCTCTCAATCAGAGTTGTTCTTCAATGAGAAGAG 410
 QY 121 IleTyrAlaProPheYserLeuLeuProMetValAsnProAspIleValIlePheGly 140
 DB 411 ATCTATAGCCCTTCAAAAGCCTTCTTCCATGGTCAACCCAGATACAGTAGTGTGGG 470
 QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValIlePheAsp 160
 DB 471 GGATGGGATCAGCAATATGAAATTTGGCTGATGCCATGGCAGGGCCAAAGGTTAGAT 530
 QY 531 ATTTGATCTCCAGAGCAGCTCAGGCCCTCATGAGACACATGGTCCCACTCCCTGGATC 590
 DB 181 TyrAspProAspPheIleAlaAlaAsnGlnIleGluArgAlaAsnValIleYserGly 200
 QY 591 TAGAGTCTGATTTTCAAGCCGATCAGGATCACTGGCCAAACGTTGATTAAGA 650
 DB 201 ThrLysGlnIleGlnValIleGlnIleIleLysAspIleLysValIlePheLysGlnIleThr 220
 QY 651 ACCAAGAAAGAACAGATTCAACATCATCANAAGATGAGGACCTTTAAGAGCAAAAT 710
 DB 221 LysValIleAspLysValIleValIleThrTrpAlaAsnThrGluArgYserAsnLeuVal 240
 QY 711 AAGGTGACAAAGTGTAGTGGCTGTCGACAGCCAAACAGAAAGATACGATATGTTGTC 770
 DB 241 ValIleLysLeuAsnAspThrMetGluAsnLeuAlaIleValIleAspArgAsnIleGlu 260
 QY 771 GTTGGGCTTAAAGATACACCCGCAACGTCGATGAGGATGAGGATGAGGCGAG 830
 DB 261 IleSerProSerThrLeuTyrAlaIleAlaCysValIleMetGluAsnValIleProPheIleAsn 280
 QY 831 AATATCTCTTCAACCTTGTATGCAATAGCTTGTGTGTAAGATGTTCCGTTCAATTAAT 890
 DB 281 GlySerProGlnAsnThrPheValIleProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
 QY 891 GGCAGCCCAAAATATCTTTGTTCAGGGGTTGATTGATTGGCGATTCAAGAGAACT 950
 DB 301 LeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMetLysSerValIleValIleAsp 320
 QY 951 TTGATTTGGGAGATGATTTTCAAGCGGTCAAAACCAAGATGAATCACTGCTAGTTGAT 1010
 DB 321 PheLeuValIleGlyAlaGlyIleLysProThrSerIleValSerTyrAsnIleGluIleAsn 340
 QY 1011 TTCTCTTGTGGAGCTGGTATCAAGCCCAAGCTGATTTGATGATATCATCTGGGAAAC 1070
 DB 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
 QY 1071 AACGACGGGATGATCTATCAGACACCCAAATCTTCGTCCTCAAGAGATTTCTTAAGAC 1130
 DB 361 AsnValIleAspAspMetValIleAsnSerAlaIleIleValIleGluProGlyGluIlePro 380
 QY 1131 AATGTGTAGACGACATGCTTCCAGCAACGCGATCTTATGAGCTGTGTAACACCCC 1190
 DB 381 AspHisValIleValIleLysTyrValIleProTyrValIleLysSerLysArgAlaMetAsp 400
 QY 1191 GACCAATATGTTGTCTCAAGATGATGACCGTATGTGGGAGACAGCAAGACCAATGTGAC 1250
 DB 401 GlyTyrThrSerGlnIlePheMetGlyGlyLysAsnThrIleValIleLeuIleAsnThrCys 420
 QY 1251 GAGTACACTTCCGAGATATTCATGGAGGGAAGAAAGTACATAGTTTGCACAAACAGTGT 1310
 DB 421 GluAspSerLeuLeuAlaIleProIleIleLysAspLeuValIleLeuAlaGluLeuSer 440
 QY 1311 GAGGACTCTCTTCTGCTGCTCCCATCATCTGAGATTTGTGCTCTTACATAACTACAGC 1370
 QY 441 ThrArgIleGlnPheLysValIleGluAsnGluGlyLysPheHisSerPheIleProValAla 460

DB 1371 ACTCGATTCACTCAAAAGCCAGAGAGGAGCAAAATTCATTCTGTTCCACCGGTGGA 1430
 QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValIleProGlyIleThrProValIleAsn 480
 DB 1431 ACCATCTTCAGCTACCTTAACCAAGCCCTTGTATACCTCCAGGACACCGGTCTGCAAC 1490
 QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValIleGluAla 500
 DB 1491 GCACTGTCAAGCAGCGTCCATGCTCGAAGACATCTTGAGGCTTGTGTGATTTGGCT 1550
 QY 501 ProGluAsnAsnMetIleLeuGluTyrIleYs 510
 DB 1551 CCAGAAACATATATGATTTTGGATATCAAG 1580
 RESULT 10
 LOCUS BT013759 1954 bp mRNA linear PLN 11-MAY-2004
 DEFINITION Lycopersicon esculentum clone 132642F, mRNA sequence.
 ACCESSION BT013759
 VERSION BT013759.1 GI:47105174
 KEYWORDS FLI_cDNA.
 SOURCE Lycopersicon esculentum (tomato)
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 1954)
 AUTHORS Kirness, B.F., Wang, W. and Vazeille, A.
 TITLES Direct Submission
 JOURNAL Submitted (11-MAY-2004) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA
 FEATURES
 source 1..1954
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /db_xref="taxon:4081"
 /clone="132642F"
 /issue_type="fruit"
 /note="TZBX41"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,226-186 Length: 1954
 Score: 2437.00 Matches: 464
 Percent Similarity: 95.49% Conservative: 23
 Best Local Similarity: 90.98% Mismatches: 23
 Query Match: 92.63% Indels: 0
 DB: 8 Gaps: 0
 US-10-718-952-12 (1-510) x BT013759 (1-1954)
 QY 1 MetPheIleGluAsnPheLysValIleGluSerProAsnValIleYserThrGluThrGluIle 20
 DB 107 ATGTTTATTGAGAACTTTTAAGTTTACAGCCCCCAATGTTAAGTACACAGAGTGAAGATT 166
 QY 21 GlnSerValIleTyrAsnTyrGluThrThrGluLeuValIleGluAsnArgAsnIleThrTyr 40
 DB 167 CACTCTGTGTATGATTTATCAAAACACTGAGTTGTTTCATGAGGAAGAAAATGAGACATTT 226
 QY 41 GlnTrpIleValIleLysProLysSerValAsnTyrGlnPheLysThrAsnThrHisValIlePro 60
 DB 227 CAATGAGCTGTAAGACCAAAACGTGCAAAATATGAAATTCAAACAGATGTTCATGTTCCA 286
 QY 61 LysLeuGlyValIleMetLeuValIleGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
 DB 287 AATATTAGAGATTATCTTGTGATGGAGTGAACCAATGTTTCAGCTTACCCGAGAGT 346
 QY 81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnIleAlaAsn 100
 DB 347 GTTATTGCTAATAGAGAGAAATTCATGGGCTACCAAGATTAAGTGCAGAGCAACCAAT 406
 QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValIleGlySerPheGlnIleGluIle 120

Db 407 TACTTGGTCTCTTACACAGGCTCTCTACTATTGAGTGGTCTTTCAATGAGAGAG 466
Qy 121 IletyralaProPhylSerLeuEupProMetValAsnProAspAspIleValPheGly 140
Db 467 ATCTATGCTCCCTTTTAAGACACTACTTCCCATGCTGCTCAATCCAGATGATGATTTTGGCA 526
Qy 141 G1YTPAspIleSerAsnMetAsnLeuValAspAlaMetAlaArgAlaIleValPheAsp 160
Db 527 GGATGGAGCATCAGCAACATGAAATTTGGCAGATGCATGCGCAAGAGCTAAAGTTTGGAC 586
Qy 161 IleaAppleGlnLysGlnLeuArgProCtyrMetGluSerMetValProLeuProGlyIle 180
Db 587 ATTTGATCTTCAAAATGCAATGAGCGGCTGACATGAGATTCATGCTCCACTGCTGATAC 646
Qy 181 TTTAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
Db 647 TATGACCTGATTTCTTCTGCTCAACCAAGGCTCAGTGTCAATGCTCAATGCTCAAGGA 706
Qy 201 ThrLysGlnGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
Db 707 ACCAAGAAAGAACAAATGATCATCAATTAATTAAGATATTAGGAGTTTAAGAGAACAAAC 766
Qy 221 LysValAspLysValValIleLeuTyrPheAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db 767 AAGGTGACAGGAGTGGTCTGCTGAGCTGCAACACTGAAAGATACAGCGTGTGGCT 826
Qy 241 ValIleGlnLeuAsnAspPheMetGluAsnLeuValAlaValAspArgAsnGlnAlaGly 260
Db 827 GTTGGCTTTAATGATACCATGAGAAACCTTCTTGTCTGCTGATGATGAAATAGGCTGAA 886
Qy 261 IleserProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 887 ATATCTCTTCCACCTTGTATGCTAATGCTTGTGATCTTGAATAAGTACCTTCAATCAT 946
Qy 281 G1YSerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 947 GGAAGGCCCAAAATGATCTTTGTCCAGGCTCATGATTTGGCCATCAAGAGAACACT 1006
Qy 301 LeuIleG1YValAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 1007 TTGATTTGGTGGTGAAGTACTTTAAGAGTGTCAAAACCAATGAAATGATGCTGGTTGAT 1066
Qy 321 PheLeuValG1YAlaG1YIleLysProThrSerIleValSerTyrAsnIleLeuGlyAsn 340
Db 1067 TTCCTTTGGTGAAGCTGGTATTAAGCCAAATCATCATGATGATCAACCATTTGGGTAAAC 1126
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
Db 1127 AATGATGGAATGAAATCTGTCTGCCCTCAACCTTCCGTTCAAGAGAGATCTCGAATAAGC 1186
Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlnIleAspPro 380
Db 1187 AACGTTTGTGAATGACATGCTTCTAGTACGCCATCTTTAATGAGCTGGAAGACATCTCT 1246
Qy 381 AspHisValValIleLysTyrValProTyrValG1YAspSerIleValArgAlaMetAsp 400
Db 1247 GACCATGTTGTTGTATTAAGTACCATATGTCGAGATAGCAAGAGGCAATGAT 1306
Qy 401 G1UTYrThrSerGlnIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys 420
Db 1307 GAGTACACATCAGACAGATTTTCAATGAGTGAAGAGAACACATTTGATTCACAATACCTGT 1366
Qy 421 G1UAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1367 GAGGATTTCTTTTGGCTGCTCCAAATTAATTTGATTTAGTCTCTGCTGATCACTCAAT 1426
Qy 441 ThrArgIleGlnPheLysAlaGluLeuGlnGlyLysPheHisSerPheHisProValAla 460
Db 1427 ACCCGATTCAGTCAAAAGCTGAAGAGAGGGTAATTCATCTCTTCATCCCGTGGGG 1486
Qy 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480

Db 1487 AACATCTCACTTATCTTACCAAGGCTCTTGGTACCAACAGGTACACCACTGATGAT 1546
Qy 481 AlaLeuSerLysGlnArgAlaMetLeuGlnLeuAlaMetAlaArgAlaCysValG1YLeuAla 500
Db 1547 GCCCTTTAAAGCAGAGGCAATGCTGCTGAACATATTAGAGGCTTGTGTTGACTGGCA 1606
Qy 501 ProGluAsnAsnMetIleLeuGlnTyrLys 510
Db 1607 CCAGAGAACAAACATGATCTTGAATATACAA 1636
RESULT 11
AF120146
LOCUS AF120146 1538 bp mRNA linear PLN 23-APR-1999
DEFINITION Triticum aestivum cultivar Fielder myo-inositol 1-phosphate
synthase (MIPS) mRNA, complete cds.
ACCESSION AF120146
VERSION AF120146.1 GI:4589061
KEYWORDS
SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 1538)
Hussain, A., Yan, W., Bock, C., Baga, M., Chibbar, R. and Georges, F.
CDNA clone for myo-inositol 1-phosphate synthase from wheat
Unpublished
2 (bases 1 to 1538)
Hussain, A., Yan, W., Bock, C., Baga, M., Chibbar, R. and Georges, F.
Direct Submission
Submitted (14-JAN-1999) National Research Council Canada, Plant
Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N
0W9, Canada
FEATURES
source
1..1538
location/Qualifiers
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Fielder"
/db_xref="taxon:4565"
1..1538
/gene="MIPS"
1..1533
/gene="MIPS"
/EC_number="5.5.1.4"
/function="converts glucose 6-phosphate into myo-inositol
1-phosphate"
/note="cytosolic enzyme"
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAD26330.1"
/db_xref="GI:4589062"
/translation="MFIESFVSES PNVKYTEINHSVYETRETVHNRNGTYQWV
KPKTVKVDTRVPLKGLVLMVGMGNGSTLTAGVYANKGISMATKDVQOANYF
GSLQASSTRGSGNGBEIVAPKSLPMNPPDIVRGMDISMNILADANARAYLD
IDLOKQLPVEHWPPURGIYDPPPIAANGSRANSYKTKKEQVQDIIDMKEPKE
KNKDKLVLTANTATERSDVVGALNDTMELLASVEKSESISPTLYALACVLEGI
PFINGSPQNTVPGLIEALISKNTLIGDDPKSQTKKSVLVDFLVGAGIKPTSYIS
YNHGNDGNMLSAPOFRSKEIKSKSNVDVMVANSGLIFEGEPDHVVYIKVPPV
GDSKRAMDEYTSEIFMGSKNTIYVHNTCEDELAAPIILDVLAEISTRLOFAEGE
GKFSHFVAITLISYTKAPLVPAGTIVNALISQRMLLENILALACVGLAPENNMLE
YK"

Alignment Scores:
Pred. No.: 8,67e-184 Length: 1538
Score: 2400.00 Matches: 456
Percent Similarity: 95.69% Conservative: 32
Best Local Similarity: 89.41% Mismatches: 22
Query Match: 91.22% Indels: 0
Gaps: 0
DB: 8
US-10-718-952-12 (1-510) x AF120146 (1-1538)

QY 1 MetPhe11eGluAsnPhelYsVal1GluSerProAsnVal1IleTyrThrGluThrGlu1 20
 Db 1 ATGTTATGAGAGGCTTCAAAAGTTGAGAGCCCAAGCTGAAGTACACAGGAACGAGATC 60
 QY 21 GlnSerVal1TyrAsnTyrGluThrGluLeuVal1H1eGluAsnArgAsnGlyThrTyr 40
 Db 61 CATTCGTGTAGCATTAACGAACCAACGAGAGTGTGTCATGAGAACGGTAAACGGCACTTAC 120
 QY 41 GlnTrr11eVal1LysProLysSerVal1AsnTyrGlnPhelYsThrAsnThi1sVal1Pro 60
 Db 121 CAATGGGTGCTCAAGCCCAAAACCGTCAATACGATTTCAAAACACACACTCCCGTCCCC 180
 QY 61 LysLeuGly1ValMetLeuVal1GlyTrrPgl1GlyAsnAsnGlySerThrLeuThrGly1 80
 Db 181 AACCTGGGCTTATGCTCGTTGATGGAGGAGGAAACATGATCAACCTTAACCGCTGGA 240
 QY 81 Val11eAl1AsnArg1LysP11eSerTrrPala1ThrLysAsnLys11eGlnGln1Asn 100
 Db 241 GTGATCGCCCAACAAAGAGGATCTGTGGCGACTTAAGACAAAGTGCAACAGCAAC 300
 QY 101 TyrPheGlySerLeuThrGln1AsnAla1IleArgVal1GlySerPheGlnGlyGlu 120
 Db 301 TACTTGGCTCCCTCACTCAAGCTTCTTCGATCCGGGTGGGATCTTAACATGGTGAAG 360
 QY 121 11eTrrAla1ProPhelYsSerLeuLeuProMetVal1AsnProAsnArg11eAla1PheGly 140
 Db 361 ATCTACGCTCTTTCAGAGACCTCTCCCATGATGTAACCTGACATGTTGTTGGT 420
 QY 141 GlyTrrPasp11eSerAsnMetAsnLeuAla1AspAla1MetAla1ArgAla1LysVal1PheAsp 160
 Db 421 GGTGGGACATTAAGCAGCATGAACCTTAAGCAGCCTATGGCTTAAGACAGAGCTCTTGAC 480
 QY 161 11eAspLeuGln1LysGlnLeuArgProTyrMetGluSerMetVal1ProLeuProGly11e 180
 Db 481 ATTGATCTCCAGAAACAGCTCAGCGCTACATGACACATGGTCCACACTCCCTGGTATC 540
 QY 181 TyrAspProAspPhe11eAla1AsnGlnGluLysArg1AsnAsnVal111eLysGly 200
 Db 541 TACGACCTGACTTCAATCGCTGGCAATCAAGGGTCAAGTCTAATAGCGTGAACAAAGG 600
 QY 201 ThrLysGlnGlnGlnVal1GlnGln11eLysAsp11eLysAla1PheLysGln1Ala1Thr 220
 Db 601 ACCAAGAAAGAAACAATCGACGACGATCATTAAGACATGAGGAGCTTAAAGGAAGAAC 660
 QY 221 LysVal1AspLysVal1Val1LeuTrrPala1AsnThrGluArgTyrSerAsnLeuVal 240
 Db 661 AAGGTGATTAAGCTGTGTTCTGTGAGACGCAACACAGAGGCTTAACAGCGATGGTGC 720
 QY 241 Val1GlyLeuAsnAspThrMetGluAsnLeuLeuAla1Val1AspArgAsnGln1Ala1Glu 260
 Db 721 GTAGGGCTTAACACAGATGAGGAATCTAATGGCTTCTGTGAAGAAACGAGTGTGAG 780
 QY 261 11eSerProSerThrLeuTrrAla1IleAla1CyVal1MetGluAsnVal1ProPhe11eAla 280
 Db 781 ATCTGCCCTTCAACGCTATATGCGATTGCTGTGTTCTTGAAGGCACTCTCTTCAATTAAT 840
 QY 281 GlySerProGlnAsnThrPheVal1ProGlyLeu11eAspLeuAla11eAla1ArgAsnThr 300
 Db 841 GGGAGTCCCAAGAACGCTTGTCTTGAGCTCATGAGCTGAGCTATATCAAGAACTGT 900
 QY 301 Leu11eGly1GlyAspAspPheLysSerGlyGlnThrLysMetLysSerVal1LeuVal1Asp 320
 Db 901 TTGATCGGTGGGATACCTCAAGAGTGTGACACATGAAGTGAAGTCTGTTTGGTGTGAT 960
 QY 321 PheLeuVal1GlyVal1Gly11eLysProThrSer11eVal1SerTyrAsnH1sLeuGly1Asn 340
 Db 961 TTCCTTGTGGAGCTGGCATCAAGCTACTGATTTGTAGACTAATCATTTGGGAAAC 1020
 QY 341 AsnAspGlyMetAsnLeuSerAla1ProGlnThrPhaArgSerLysGlu11eSerLysSer 360
 Db 1021 AATGACGGGATGAACCTCTCAGCGCTCAAGACTTTTATGATCTTAAGAGATCTCAAGAGAC 1080
 QY 361 AsnVal1Val1AspAspMetVal1AsnSerAsnAla11eLeuTrrPala1ProGlyGlnH1sPro 380

Db 1081 AATGTTGTTGACACATGTTGCTTACTAACGATTCCTCTTTGAGCTGTGAACATCTCT 1140
 QY 381 AspH1sVal1Val11eLysTrrVal1ProTrrVal1GlyAspSerLysArgAla1MetAsp 400
 Db 1141 GACCATGTTGTGTCATCAAGTATGTTCTTATGTTGGGATATGTAAGAGACCATGAGAC 1200
 QY 401 GlnTrrThrSerGln1LysPheMetGlyGlyLysAsnThr11eVal1LeuH1sAsnTrrYs 420
 Db 1201 GAGTACACCTCGGAGATATTCATGGAGGGAAGAAATACATGTTATGATTAACACTTGT 1260
 QY 421 GluAspSerLeuLeuAla1Pro11eLysAspLeuVal1LeuLeuVal1GluLeuSer 440
 Db 1261 GAGATTCCTCTTTGGCTGCTCCATCATCTTGATCTTGTCTTCCGTGAGCTCAAC 1320
 QY 441 ThrArg11eGlnPhelYsAla1GluAsnGlnGlyLysPheH1sSerPheH1sProVal1Ala 460
 Db 1321 ACCAGAAATCAATTCACAACTGAAGGAGGAGGAGGATTCATTTCTTCCACCAGTAGCA 1380
 QY 461 Thr11eLysSerTrrLeuThrLysAla1ProLeuVal1ProProGlyThrProVal1Val1Asn 480
 Db 1381 ACCATATCTAGTTACCTCAAAAGGACCTCTTGTACAGCCGGGACACCTGTGTGAAC 1440
 QY 481 AlaLeuSerLysGlnArgAla1MetLeuGluAsn11eMetArgAla1CyVal1GlyLeuAla 500
 Db 1441 GCTTATCTTAAGCAGAGGCTATGCTCGAGAAACATCTCAAGGCTTGCGCTTGGC 1500
 QY 501 ProGluAsnAsnMet11eLysGlnTrrLys 510
 Db 1501 CCAGAGAACACATATCTTGAATATCAAG 1530

RESULT 12
 AF120147
 LOCUS 1538 bp mRNA linear PLN 23-APR-1999
 DEFINITION Triticum aestivum cultivar Taber myo-inositol 1-phosphate synthase (MIPS) mRNA, complete cds.
 ACCESSION AF120147
 VERSION AF120147.1 GI:4589063
 KEYWORDS
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.
 1 (bases 1 to 1538)
 Husain,A., Yan,W., Bock,C., Baga,M., Chibbar,R. and Georges,F.
 cDNA clone for myo-inositol 1-phosphate synthase from wheat
 Unpublished
 2 (bases 1 to 1538)
 Husain,A., Yan,W., Bock,C., Baga,M., Chibbar,R. and Georges,F.
 Direct Submision
 Submitted (14-JAN-1999) National Research Council Canada, Plant
 Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N
 0W9, Canada

FEATURES
 source
 1..1538
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Taber"
 /db_xref="taxon:4565"
 1..1538
 /gene="MIPS"
 1..1533
 /gene="MIPS"
 /EC_number="5.5.1.4"
 /function="converts glucose 6-phosphate into myo-inositol 1-phosphate"
 /note="cytosolic enzyme"
 /codon_start=1
 /product="myo-inositol 1-phosphate synthase"
 /protein_id="AAD26331.1"
 /db_xref="GI:4589064"
 /translation="MFISFKVESPNVKTENEHSYDYETTEVVENRNTQMVV"

KKRTVKYDFKTDTRVPKLGVLVGVGGNNSTLTGVIANKEGISMATKXVQOANYF
 GSLTQASSIRVGSYNGEIEIYAFKSLIPVNPDDVDFGMDISDMNLADIKMRFEKE
 IDLOLRAPRYEMHWYIPGCIYDPDPAANGDRANSVIGKTEOVDOIDIKMRFEKE
 KRKVDKIVLTMTANTERSYDVVGVNDMENTLASEVENKESIPTSTVIALACVGI
 PRINSNPQTFYPGILBELAISXKCLIGDDFKSGOTKMSVIVDLVNGAGIRPISVS
 YNHLNNDNDMNLAPOTERKESIKSNVDDVANSNGILFEGEHPDHVVKIYVYV
 GDSKAMBEYTSIEFMGKNITVMNTECDSLAAPIILDLVLELSTRLOPKRGE
 GKFSFHPVATILSYLTAKAPVAGTPPVNALSKORAMLENTLRACVGLPENNMLE
 YK"

ORIGIN

Alignment Scores:
 Pred. No.: 8, 67e-184 Length: 1538
 Score: 2400.00 Matches: 456
 Percent Similarity: 95.69% Conservative: 32
 Best Local Similarity: 89.41% Mismatches: 22
 Query Match: 91.22% Indels: 0
 DB: 8 Gaps: 0

US-10-718-952-12 (1-510) x AF120147 (1-1538)

QY 1 MetpHeileGluAsnPhelYserProaenValYserThrgluThrgluile 20
 Db 1 ATGTTTCATGAGAGCTTCAAAGTTGAGACCCAAACGTGAAGTACACAGAGAACGAGATC 60
 QY 21 GlnSerValTYrAsnTYrGluThrThrgluLeuValHiegluAsnArgasnGlyThrTYr 40
 Db 61 CATTCGTCTGATGATTACGAAACACGAGAGTGTCTCATGAGAACCGTAAACGGCACTTAC 120
 QY 41 GlnTPrIleValYserProLYserValAsnTYrGlnPhelYserThraSnThrhieValPro 60
 Db 121 CATGGGTGCTCAAGCCCAAAACCGTCAATATGATTCATAAAGACAGACCTGCGCTCCC 180
 QY 61 LyeleuGluValMetleuValGlyTPrGlyYAsnAsnGlySerThreuthrGlyGly 80
 Db 181 AAGCTAGGCGTATGCTCGTTGATGGAGGAGAAACAATGATCAACCTTAAACCGCTGGA 240
 QY 81 ValIleAlaAsnArgGluAspIleSerTPRAlaThrLYAspLYseIleGlnGlnAlaen 100
 Db 241 GGTATGGCCAAACAAGAGGATCTGTGGCGACCTAAAGACAAGTGCACAAGCAAGAAC 300
 QY 101 TYrPheGlySerleuThrglnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
 Db 301 TACTTGGCTCCCTCACTCAAGCTTCTTGATCCGGGTGGATCTTACATAGTGTGAAGAG 360
 QY 121 IleyrAlaProPhelYserleuProMetValAsnProAspIleValPheGly 140
 Db 361 ATCTACGCTCTTTCAAAGAGCTCTCCCATGAGTAAACCTGACGATGTGTGTGGT 420
 QY 141 GLYTPRAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValPheAsp 160
 Db 421 GGTGGGACATTAAGCAATGAACTTACGACAGCGCTATGGCTAGCAAGAGCTCTGAGC 480
 QY 161 IleAspLeuGlnYserGlnLeuArgProTYrMetGluSerMetValProleuProGlyIle 180
 Db 481 ATTGATCTCCAGAAACAGCTCAGCGCTACATGAGCAACATGGTCCACTCCCTGATATC 540
 QY 181 TYrAspProAspPheIleAlaIleAsnGlnGluArgAlaAsnAsnAlIleyGly 200
 Db 541 TACGAGCCCTGACTTCCTGCTGCAATCAAGGCTCAGTGTATATGCGTATCAAAAGGT 600
 QY 201 ThrLYserGlnGlnAlaGlnIleIleIleYAspIleYAsnAlaPheYserGluAlaThr 220
 Db 601 ACCCAAAAGAAACAATCGACCAACGATCATCAAAAGACATAGAGAGTTTAAAGAGAAAGAC 660
 QY 221 LysValAspLYserValValleuTYrThrAlaAsnThrGluArgTYrSerAsnLeuVal 240
 Db 661 AAGGTGATTAAGCTGTGTCTGTGAGACAGCCAAACAGAGGGTTACAGCGATGTGTC 720
 QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaIleValAspArgasnGluAlaGlu 260
 Db 721 GTAGGGCTTAAACGACGATGAGGAATCTTATGCTGTGAGAAAGACGATCTGAG 780

QY 261 IleSerProSerThrleuTYrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
 Db 761 ATCTGCGCTTCAACGCTATGCGATTCCTGTGTTCTTGAAGGATCTCTTCAATTAAT 840
 QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgasnThr 300
 Db 841 GGAATCCGCGAACAACCTTTGTTCTGACTCATCGAGCTGGCTGATATCGAAGACGTGT 900
 QY 301 LeuIleGlyGlyAspAspPheLYserGlyGlnThrLYserMetLYserValleuValAsp 320
 Db 901 TTGATCGGTGGGAGATGACTTCAAGAGTGTGTGACATAAGATGAATCTGTTTGGTTGAT 960
 QY 321 PheleuValGlyAlaGlyIleLYserProThrSerIleValSerTYrAsnHieGluAsn 340
 Db 961 TTCCTTTCGAGAGCTGGCATCAAGCCCTCTTCGATTTGGAGCTAATATCATTTGGGAAC 1020
 QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLYserGluIleSerLYser 360
 Db 1021 AATGACGCGATGAACCTTCAGCGCTCAAGCTTTAATCTTAAAGATTCGAAGAGAC 1080
 QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTYrGluProGlyGluHiePro 380
 Db 1081 AATGTTGTGACGACATGTGGCTATGACGTATCCCTTTGAGCTGGTGAACATCTCT 1140
 QY 381 AspHieValValIleLYserTYrValProTYrValGlyAspSerLYserAlaMetAsp 400
 Db 1141 GACCATGTTGTGCTCATCAAGATATGTTCTTATGTTGGGATGATGAAGACCATGAGAC 1200
 QY 401 GlnTYrThrSerGlnIlePheMetGlyLYserAsnThrIleValLeuHieAsnThrCys 420
 Db 1201 GAGTACACCTCGAAGATATTCATGGAGGAGGAAGATGATTTATCATCAATACCTTGT 1260
 QY 421 GluAspSerleuLeuAlaIleProIleIleLeuAspLeuValleuLeuAlaGluLeuSer 440
 Db 1261 GAGATTTCTCTTTTGGCTGCTCCATCATCTTGATCTTGTCTTCTGCTAGCTGACG 1320
 QY 441 ThrArgIleGlnPhelYserAlaGluAsnGlyLYserPheHieSerPheHieProValAla 460
 Db 1321 ACCAAGATTCATTAATCAAAGCTGAAGGGAGGGAGGATTCATTTCTTCCACCACATACA 1380
 QY 461 ThrIleLeuSerTYrleuThrLYserAlaProleuValProProGlyThrProValAlaAsn 480
 Db 1381 ACCAATACCAAGTATCACTCAAAAGGACCTCTTGAACAAGCCGGGACCTGTGTGTAAC 1440
 QY 481 AlaLeuSerLYserGlnArgAlaMetleuGluAsnIleMetArgAlaCysValGlyLeuAla 500
 Db 1441 GCTTATCTTAAGACAGAGGCTATGCTGAGAAACATCTCAGGCGCTGCGTTGGCTTGG 1500
 QY 501 ProGluAsnMetIleleuGluTYrLYs 510
 Db 1501 CCAGAGAACCAATGATCTTGAATACAAAG 1530

RESULT 13
 AF120148 1538 bp mRNA linear PLN 23-APR-1999
 LOCUS Triticum aestivum cultivar Biggar myo-inositol 1-phosphate synthase
 DEFINITION (MPS) mRNA, complete cds.
 ACCESSION AF120148
 VERSION AF120148.1 GI:4589065
 KEYWORDS
 SOURCE
 ORGANISM
 Triticum aestivum (bread wheat)
 Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.
 1 (bases 1 to 1538)
 Huseain,A., Yan,W., Bock,C., Baga,M., Chibbar,R. and Georges,F.
 cDNA clone for myo-inositol 1-phosphate synthase from wheat
 JOURNAL
 Unpublished
 2 (bases 1 to 1538)
 Huseain,A., Yan,W., Bock,C., Baga,M., Chibbar,R. and Georges,F.
 Direct Submission
 TITLE
 Submitted (14-JAN-1999) National Research Council Canada, Plant

Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Canada

FEATURES

Location/Qualifiers

Source

1. .1538

```
/organism="Triticum aestivum"
```

```
/mol_cype="lucwa"
/cultivar="Biggar"
```

/db_xref="taxon:4565"

gene

1. .1538

CDS

```
1. gene="MLPS"
1. 1533
```

```
/gene="MIPS"
```

```
/EC_number="5.5.1.4"
```

```
/function="converts gluc
1-phosphate"
```

```

- product
/note="cytosolic enzyme"

```

```
/codon_start=1
```

```
/product="myo-inositol 1
/protein id="AAD26332.1"
```

```
/proc/self/exe=/usr/bin/perl
```

```
/translation="MFIESFKVESPNKYTENEIHSVYDYETTEVHNRNGTYQWV
```

CSIT0A86SIBWCCVNGEETIYADREKCI DUDUDDNDU/EGGWDGCVUT ADUWABADUUF
KPKTVKVBDFKIDTRVPKLGVMVLVWGCGNNGSTLTAGVIANKEGISWATKDKVQQANYF

1DLOKOLBYMEHMYPLPGTYDPDEIAANGSBANSYTKTKKEOYDOTKDMREEKEE
 35B1QASSIANGSEETAPFNSLEPMVNPDDVFGWDISDMINLALAMAKARVLDD

KNKVDKLVLWTANTERYSDVVGLNDTMENLLASVEKNESEISPSTLYAIAACVLEGI

PFINGSPÖNTFVPGLIETLAI SKNCLIGDDFKSGÖTKMKSVLDFLVGAGIKPTSI VSS

GD\$K\$BAMDEYTSIEMG\$KNTIMHNTCE\$DI.I.A\$PII.I\$M.I.A\$E\$.STBIOEKAEGEE\$
INHLGNNDGML\$APQIFR\$K\$EISK\$NVDDMVASNGILFE\$E\$G\$EHPDHVVILKVPYV

GKEHSHFHPVATILSYLTKAPLVAGTPVNVNLSKORAMLENILRACVGLAPENNMILLEE

$$YK''$$

ORIGIN

Alignment Scores:

Pred. No.: 8.67e-184

Score: 2400.00
Percent Similarity: 95.60%

Best Local Similarity: 89.41%

Query Match: 91.22%

DB: 8

US-10-718-952-12 (1-510) X AFI20148 (1-1538)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1 MetPheIleGluAsnPhelysValGluSer

1. $\Delta T_{\text{max}} = 10^\circ\text{C}$

LOCUS	AF433879	1986 bp	mRNA	linear	PLN 07-NOV-2001
DEFINITION	AF433879	Suaeda maritima subsp. salsa myo-inositol-1-phosphate synthase (INPS) mRNA, complete cds.			
ACCESSION	AF433879				
VERSION	AF433879.1	GI:16755888			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
gene					
CDS					
Alignment Scores:					
Pred. No.:	1,436-183	Length:	1986		
Score:	2399.00	Matches:	462		
Percent Similarity:	94.72%	Conservative:	22		
Best Local Similarity:	90.41%	Mismatches:	25		
Query Match:	91.18%	Indels:	2		
DB:	8	Gaps:	2		
US-10-718-952-12 (1-510) x AF433879 (1-1986)					
1 MetPheIleGluAsnPhelYsValGluSerProAsnValYsYrThrGluThrGluIle	20				
140 ATGTCATCGAAAATTTTCAGGTAGAAAGCCCAAAATGTTAACTACAGAAATCGAAATTT	199				
21 GlnSerValYrAsnTyrGluThrThrGluLeuValIhsgluAsnArGAsnGlyThr---	39				
200 CATTGAGTTTCAATTAAGAAACCACTAAATTAAGTTCCAGGAATCGCAAAAGGCTTGCT	259				
40 TyrGlnTrpIleValYsProYsSerValAsnTyrGlnPheYsThrAsnThrHisVal	59				
260 TATCATGAGATGCTAAACCTTAAACCGCTCAATTCCAATTCCTAAACCTCAACATCATATT	319				
60 ProYsLeuGlyValMetLeuValGlyTrpGlyYsAsnAsnGlySerThrIleuThrGly	79				
320 CCAAACTTGGGGTGATGTTGGTGGAATGGGGTGCAACAACGGTTGCACACTACCGGT	379				
80 GlyValIleValAsnArGluAspIleSerTrpAlaThrIlyAspIlyIleGlnGlnAla	99				

D5	380	GGTGTATCGCAATCGCAGGGGATATCATGGGCAACAAGACAATAATTCAACAAGCT	439
OY	100	AsnTyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu	119
D5	440	AACATATTTGGATGCCCTCACCTCAGGGCTTCTTCAATTCGTGTGGTCTTTAATGGAG	499
OY	120	GlnIleTyrAlaProPheGlySerLeuLeuProMetValAsnProAspAlaIleValPhe	139
D5	500	GAGATTATATGCTCTTTTAAAGAGCTTCTTCTTAATGTGTGAACCAATGATATATGTCTT	559
OY	140	GlyGlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaValPhe	159
D5	560	GGGGGCTGGGACATTAAGTGAACCTTGACAGAGCCATGACGAGGGCAAGGTGTTG	619
OY	160	AspIleAspLeuGlnValGlnLeuArgProTyrMetGlySerMetValProLeuProGly	179
D5	620	GACATTAATTTGGCAAAAGCAGTTGAGACCTTAAATGAGACATGTGTTCCCTTCCTGT	679
OY	180	IleTyrAspProAspPheIleAlaAlaAsnGlnGlnLysArgAlaAsnAsnValIleLeys	199
D5	680	ATCTATAGACCTGTATTTTCATTTGCTGCTAACCAAGATTTCTGTGCAATACATCATTAAG	739
OY	200	GlyTyrHisGlnGlnGlnValGlnGlnIleIleLeysAspIleLeysAlaPheLysGluAla	219
D5	740	GGCACTAAGAAAGGACGAAGTTGAGCAAGCATTAAGGACATCAGGGAATTTAAGAGAAC	799
OY	220	ThrLysValAspLysValValIleLeuTyrThrAlaAsnThrGluArgLysSerAsnLeu	239
D5	800	AACAAAGTGGCAAGAGTGGTGTACTCTGGACAGCAAAACAGAGAGTCAAGTAGTGG	859
OY	240	ValValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAlaAspArgAsnGluAla	259
D5	860	GTGTGTGAGACTCAATGACACACAGAAACCTTAATGCACTGTTGAAGGAATGACTCT	919
OY	260	GluIleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIle	279
D5	920	GAGATTCTCATCGACTTGTGTGATGCAATGCTTGATCCAGAGAAATGTCCTTTTCATC	979
OY	280	AsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsn	299
D5	980	AACGATGATCCTCAGAACACCTTGTCTCCGTGTCTCATTAAGCGCCATTAAGAACAT	1039
OY	300	ThrLeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValIleVal	319
D5	1040	TGTTTGTATGGAGAGATGACTTCMAAGAGTGTGACTAAGATGAATACATGCTCTGTT	1099
OY	320	AspPheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGly	339
D5	1100	GACTTCTCTGTGGGGCTGTGATTAACCAACATCATTTGAGCTACCAACCACTTGGGA	1159
OY	340	AsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLys	359
D5	1160	AACATGATGGGATGAACCTTTCAGCACCTCAGACCTTCAGGCTCTAAAGATTTCAAG	1219
OY	360	SerAsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHis	379
D5	1220	AGTATATGTTGTATGACATGTTGCTACGAACGATCTCTGTATGAGCCTGTGGTGAAC	1279
OY	380	ProAspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMet	399
D5	1280	CCTATCATGTGTTGTTCATCAAGTATGTTCCATATTCGAGACAGCAAGAGACTATG	1339
OY	400	AspGluTyrThrSerGluIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThr	419
D5	1340	GATGAGTACACATCTGAGATTAATTCATGGGTGTGAACAACCATCGTTTGGACCAAC	1399
OY	420	CysGluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeu	439
D5	1400	TGTAGAGACTCTCTCTTGGCGCTCCATATCTCTAAGACTTGCTCACTTGGCGAAGCTT	1459
OY	440	SerThrArgIleGlnPheLysAlaGluAsnGlnGlyLysPheHisSerPheHisProVal	459
D5	1460	AGCACCCGACATTAATTCAAAGCGGAACTGAGGACAAGTTCCACACCTTCCATCCAGTG	1519

QY 460 [Methylobacterium thermophilum](#) ATCC 35061 479
 1520 [GCTACCACTACGATGATCTAATCAAGACACTCTGGTTCCACGAGTACCGGTGGTG](#) 1579
 QY 480 [Asanaleuserylgina](#) [Ala](#) [Mec](#) [Leu](#) [Glu](#) [Ser](#) [Ile](#) [Met](#) [Arg](#) [Ala](#) [Cys](#) [Val](#) [Gly](#) [Leu](#) 499
 1580 [AATGGCGTGTCAAGCAAGCAGGCGCATGTTAGAGAA---ATGAGGCGCATGTGTAGTTTA](#) 1636
 Db 500 [Ala](#) [Pro](#) [Glu](#) [Asn](#) [Asn](#) [Met](#) [Ile](#) [Leu](#) [Glu](#) [Tyr](#) [Ile](#) [Lys](#) 510
 1637 [GCTCCAGAGAACACATGATCATCTTGAGTCAAG](#) 1669
 RESULT 15.
 LOCUS MCU32511 2053 bp mRNA linear PLN 11-JUN-1996
 DEFINITION Mesembryanthemum crystallinum myo-inositol-1-phosphate synthase
 mRNA, complete cds.
 ACCESSION U32511
 VERSION U32511.1 GI:975887
 KEYWORDS
 SOURCE Mesembryanthemum crystallinum (common iceplant)
 ORGANISM Mesembryanthemum crystallinum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
 Caryophyllales; Aizoaceae; Mesembryanthemum.
 REFERENCE 1 (bases 1 to 2053)
 Ishitani,M., Majumder,A.L., Bornhouser,A., Michalowski,C.B.,
 Jensen,R.G. and Bohnert,H.J.
 Coordinate transcriptional induction of myo-inositol metabolism
 during environmental stress
 Plant J. 9 (4), 537-548 (1996)
 8624516
 2 (bases 1 to 2053)
 Ishitani,M., Majumder,A.L., Bornhouser,A., Michalowski,C.B.,
 Jensen,R.G. and Bohnert,H.J.
 Direct Submission
 Submitted (27-JUN-1995) Hans J. Bohnert, Biochemistry, University
 of Arizona, Biochemistry West 516, Tucson, AZ 85721, USA
 FEATURES
 source
 1..2053
 /organism="Mesembryanthemum crystallinum"
 /mol_type="mRNA"
 /db_xref="taxon:3544"
 182..1720
 /codon_start=1
 /product="myo-inositol-1-phosphate synthase"
 /protein_id="AAB03687.1"
 /db_xref="GI:975888"
 /translation="MFIESPVFESPNKYTENIESVYNYDTTELVAHNRKDGCGYOW
 IVRKTYOYHKHTDTRVPKLGVMVYCGNGNGSTLTGVLANRGGISMAWKDKIQAN
 YFGLTQASSTIRVGSFNGEETIYAFKSLIPVNVDDVAFGWDISDNLADAMTRARV
 FDILOQRLPYMEHWPLPGITPDPTIAQGRANNVKGTKEQVEVVIDIRF
 KERKVKVAVVLTGNTGRYSNVVAGNDTMENTIALAKNESISPTSYALACIEE
 NIPINSGPQTFVPGKIDILAKKNSIGGDFSGGQKKMSVVDPLVAGIKPTSI
 VSYVNLGNDGMSLAPQTPRSKESISNVVDVAANGILYBEGEHPDHYVYKVP
 YVNGSKRAMDEYTFSEFMCGNTIYMNATGDSLLAPPIIDVYLALESTRIOKAE
 EEDKFSHPHVAITLSTITKAPLVPQGPVYNALSKORAMENILRACVGLAPENNTI
 LEYR"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,13e-183 Length: 2053
 Score: 2395.00 Matches: 458
 Percent Similarity: 94.73% Conservative: 27
 Best Local Similarity: 89.45% Mismatches: 25
 Query Match: 91.03% Indels: 2
 Gaps: 8
 1
 US-10-718-952-12 (1-510) x MCU32511 (1-2053)
 QY 1 [Methylobacterium thermophilum](#) ATCC 35061 479

Db 132 ATGTTTATTGAGAGCTTCGAAAGTACAGAGCCCAACGTAAGCTACACAGAGATATGATT 241

QY 21 GlnSerVal¹YrAsn²Trg³Ln⁴Thr⁵Trg⁶Leu⁷Val⁸Ih⁹Isg¹⁰Uas¹¹Ar¹²Asn¹³-----Gly 38

Db 242 GAATGCGGTGACAACTACGACACCAAGGTTGGTTGATGAGAACTCGCAAGATGCTGGT 301

QY 39 Thr¹Trg²Int³Trp⁴Ile⁵Val⁶Ly⁷Pro⁸Ly⁹Ser¹⁰Val¹¹Asn¹²Trg¹³Int¹⁴Ph¹⁵Ly¹⁶Thr¹⁷Asn¹⁸Th¹⁹His 58

Db 302 GGTATTCATGATGATGATTAAAGCCCAAGACCGTTTCAGTATTCATTCGAGACTGACCTCGT 361

QY 59 Val¹Pro²Ly³Leu⁴Gly⁵Val⁶Met⁷Leu⁸Val⁹Gly¹⁰Trp¹¹Gly¹²Yas¹³Asn¹⁴Gly¹⁵Ser¹⁶Thr¹⁷Leu¹⁸Tr 78

Db 362 GTCCCAAACTCGGGGTATGCTGGTGTGATGGGGTGAATAATACAGGCTCCACCTCA 421

QY 79 Gly¹Gly²Val³Ile⁴Ala⁵Asn⁶Arg⁷Leu⁸Asp⁹Ile¹⁰Ser¹¹Trp¹²Ala¹³Thr¹⁴Ly¹⁵Asp¹⁶Ly¹⁷Ile¹⁸Gln¹⁹ 98

Db 422 GGTGGTGTGATCGCCAAACCGCAGGGAATCTTGGGCAACCAAGACAAAGATCCAGCA 481

QY 99 Ala¹Asn²Trg³Ph⁴Ly⁵Ser⁶Leu⁷Trg⁸Ala⁹Asn¹⁰Ala¹¹Iea¹²Arg¹³Val¹⁴Gly¹⁵Ser¹⁶Ph¹⁷Gly¹⁸ 118

Db 482 GCGAATCTTTTGGTTCGTTGACCCAGGCTTCCTCAATTCGGGTGGGCTTTTAAATGGA 541

QY 119 Glu¹Glu²Ile³Trp⁴Ala⁵Pro⁶Ph⁷Ly⁸Ser⁹Leu¹⁰Pro¹¹Met¹²Val¹³Asn¹⁴Pro¹⁵Asp¹⁶Ile¹⁷Val¹⁸ 138

Db 542 GAGGAGATTTATGCTCTTCCTTAAGAGCTTGCTCCCTAATGAGAACCCAGATGACCTAGG 601

QY 139 Ph¹eg²Gly³Trp⁴Asp⁵Ile⁶Ser⁷Asn⁸Met⁹Asn¹⁰Leu¹¹Ia¹²Asp¹³Ala¹⁴Met¹⁵Ala¹⁶Arg¹⁷Ala¹⁸Val¹⁹ 158

Db 602 TTTGGGGGTTGGGACATAGGACATGAACTGCTGCTGATGCAATGACCAAGGCTAGGGT 661

QY 159 Phe¹Asp²Ile³Asp⁴Leu⁵Gln⁶Ly⁷Ile⁸Val⁹Arg¹⁰Pro¹¹Trg¹²Met¹³Gly¹⁴Ser¹⁵Met¹⁶Val¹⁷Pro¹⁸Leu¹⁹Pro 178

Db 682 TTTGACATTTGATCTTCAAAAGACGCTCAGGCTCTTACATGAGACATGATGCTCCCTCC 721

QY 179 Gly¹Ile²Trp³Asp⁴Pro⁵Asp⁶Phe⁷Ile⁸Ala⁹Ia¹⁰Asn¹¹Gln¹²Glu¹³Glu¹⁴Arg¹⁵Ala¹⁶Asn¹⁷Val¹⁸Ile 198

Db 722 GGATTTTACGACCCCTGATTTATCGCTCGAACCAGGCTCCCGGCCAACAACTCAT 781

QY 199 Lys¹Gly²Trp³Ly⁴Sgln⁵Gln⁶IVal⁷Igln⁸Ile⁹Ile¹⁰Ly¹¹Ser¹²Asp¹³Ile¹⁴Val¹⁵Ala¹⁶Ph¹⁷Ly¹⁸Gln¹⁹ 218

Db 782 AAGGACACCAAGAGAGACCAAGTTGAGGAGCTATTAAAGCATTTAGGAGTTCAAGAG 841

QY 219 Ala¹Thr²Ly³Val⁴Asp⁵Ly⁶Val⁷Val⁸Val⁹Leu¹⁰Trp¹¹Thr¹²Ala¹³Asn¹⁴Thr¹⁵Glu¹⁶Arg¹⁷Ser¹⁸Asn 238

Db 842 AAGAACAAGTGGACCAAGGTGGTGGTCTTAAGACTGGCAACACAGAGAGGTACGCAT 901

QY 239 Leu¹Val²Ile³Ly⁴Leu⁵Asn⁶Asp⁷Thr⁸Met⁹Glu¹⁰Asn¹¹Leu¹²Leu¹³Ala¹⁴Val¹⁵Asp¹⁶Arg¹⁷Asn¹⁸Glu 258

Db 902 GTGGTGGTCTCGGGCTTCACAGACACATGAGAAACGTGGATCTCTGAGAAACAAACGAG 961

QY 259 Ala¹Gly²Ile³Ser⁴Pro⁵Ser⁶Thr⁷Leu⁸Trp⁹Ala¹⁰Ile¹¹Ala¹²Cys¹³Val¹⁴Met¹⁵Glu¹⁶Asn¹⁷Val¹⁸Pro¹⁹He 278

Db 962 TCGGAGATTTCTCATCCACTTCGATGCTTCGATGATGATGAGAAATATTCCTTC 1021

QY 279 Ile¹Asn²Gly³Ser⁴Pro⁵Gln⁶Asn⁷Thr⁸Ph⁹Val¹⁰Pro¹¹Gly¹²Leu¹³Ile¹⁴Asp¹⁵Leu¹⁶Ala¹⁷Ile¹⁸Ala¹⁹Arg 298

Db 1022 ATCATAGTGTACCTCCAGAACACCTTTGTCCAGGGCTCATTTGATCTGGCATCAAGAG 1081

QY 299 Asn¹Thr²Ile³Ile⁴Gly⁵Ly⁶Asp⁷Asp⁸Phe⁹Ly¹⁰Ser¹¹Gly¹²Int¹³Trp¹⁴Met¹⁵Ly¹⁶Ser¹⁷Val¹⁸Leu 318

Db 1082 AACGTTGATTTGGGTGTGATGACTTCACAGAGTGTGACCAAGATGAAATCTGTCTCT 1141

QY 319 Val¹Asp²Phe³Leu⁴Val⁵Gly⁶Ala⁷Gly⁸Ile⁹Ly¹⁰Pro¹¹Thr¹²Ser¹³Ile¹⁴Val¹⁵Ser¹⁶Trp¹⁷Asn¹⁸His¹⁹Leu 338

Db 1142 GTGACATTCCTCGTGGGCTGTGATTTAAGCCCAATCCATCAATGATGAGTACCAACCAT 1201

QY 339 Gly¹Asn²Asn³Asp⁴Gly⁵Met⁶Asn⁷Leu⁸Ser⁹Ala¹⁰Pro¹¹Gln¹²Thr¹³Ph¹⁴Arg¹⁵Ser¹⁶Ly¹⁷Glu¹⁸Ile¹⁹Ser 358

Db 1202 GCGAACATATATGGGATGAACTGTGACACCTCAAACTTCAGGTGCCAAGAGATTTCT 1261

QY 359 Lys¹Ser²Asn³Val⁴Asp⁵Asp⁶Met⁷Val⁸Asn⁹Ser¹⁰Ala¹¹Ile¹²Leu¹³Trg¹⁴Pro¹⁵Gly¹⁶Glu 378

Db 1262 AAAGCAACGTGTCATGACATGGTTGCCAGCAACGGCATCCTCTATGAGCCGGGCGAA 1321
Qy 379 HisProAspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAla 398
Db 1322 CACCCCTGATCATGTGTGTGTCATCAAGTATGTGCCATATGTGGAGACAGCAAGAGAGCC 1381
Qy 399 MetAspGluTyrThrSerGluIlePheMetGlyGlyLysAsnThrIleValIleuHisAsn 418
Db 1382 ATGGATGAGTACACTTCGGAATAATTATGGGGGGAAACGACACTATGTCTCATGCACAAC 1441
Qy 419 ThrCysGluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleuLeuAlaGlu 438
Db 1442 ACCGTGAGGACTCTCTCTTGGCTGCCCCCATCATCTGGACTTGGTCTCTCTGCTGAA 1501
Qy 439 LeuSerThrArgIleGlnPheLysAlaGluAsnGluGlyLysPheHisSerPheHisPro 458
Db 1502 CTAAAGCACCCGATTCAGCTGAAGGCTGAAGGAGAGACAACTCCACTCCATCCG 1561
Qy 459 ValAlaThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProVal 478
Db 1562 GTGGCTACCATCTGTGTTACCTCAACAGGCCCTCTTGTTCACACAGGCACCTCCAGTG 1621
Qy 479 ValAsnAlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGly 498
Db 1622 GTAAATGCACCTTCGAAGCGAGGGCTATGCTGAGAACATATGAGGGCTGTGTGGT 1681
Qy 499 LeuAlaProGluAsnAsnMetIleLeuGluTyrLys 510
Db 1682 TTGGCTCCTGAGACACATGATCCTCGAATACAAAG 1717

Search completed: June 7, 2005, 20:45:40
Job time : 5355 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 7, 2005, 17:10:41 ; Search time 633.5 Seconds

(without alignments)
4765.695 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631
Sequence: 1 MFIEFKYSPVVKXTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=ifame+.p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10718952/runcat.06062005.173400.12840/app.query.fasta.1.1294
-DB=N Geneseq.16Dec04 -QFMT=fasta -SUFFIX=ep2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=humand40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=prc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10718952@cgn.1.1.886@runat.06062005.173400.12840 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOBVERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq.16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2631	100.0	1533	12	ADQ14500 Mutant so
2	2631	100.0	1533	13	AD882001 Soybean m
3	2624	99.7	1533	12	ADQ14504 Wild type
4	2624	99.7	1533	12	ADQ14502 Mutant so
5	2624	99.7	1533	13	AD882005 Soybean m

6	2624	99.7	1533	13	AD882003	AD882003 Soybean m
7	2593	98.6	1533	12	ADQ14498	Adq14498 Mutant so
8	2593	98.6	1533	13	AD881999	Ad881999 Soybean m
9	2593	98.6	1533	13	AD881993	Ad881993 Soybean m
10	2593	98.6	1760	12	ADQ14490	Adq14490 Wild type
11	2593	98.6	1782	2	AAV62440	AAV62440 Soybean w
12	2588	98.4	1533	12	AAV62443	AAV62443 Soybean m
13	2588	98.4	1533	12	ADQ14494	Adq14494 Mutant so
14	2588	98.4	1533	13	AD881997	Ad881997 Soybean m
15	2470	93.9	1950	2	AAx90402	AAx90402 Nicotiana
16	2384	90.6	1533	6	AB213633	Ab213633 Arabidops
17	2384	90.6	1533	12	ADN73524	Adn73524 Thale cre
18	2384	90.6	1938	3	AAc50242	AAc50242 Arabidops
19	2376	90.3	1781	4	AAc87643	AAc87643 Brassica
20	2354	89.5	1665	2	AAx09006	AAx09006 Inducible
21	2349.5	89.3	1536	3	AAc43428	AAc43428 Arabidops
22	2349.5	89.3	1837	3	AAc34806	AAc34806 Arabidops
23	2346	89.2	1931	2	AAx24407	AAx24407 Maize myo
24	2337	88.8	1959	4	AAc85922	AAc85922 MIP synth
25	2285	86.8	1759	6	ABQ72653	Abq72653 Human MHD
26	2126.5	80.8	1719	3	AAc51095	AAc51095 Arabidops
27	2040	77.5	1536	12	ADP43918	Adp43918 P. coarct
28	1823	69.3	3546	2	AAx24411	AAx24411 Maize myo
29	1822	69.3	3546	2	AAx24410	AAx24410 Maize myo
30	1616	61.4	1772	4	AAH14171	AAH14171 Human cDN
31	1616	61.4	1818	4	AAH15251	AAH15251 Human cDN
32	1616	61.4	1825	6	ABT07176	ABT07176 Human ova
33	1616	61.4	1852	4	AAH47741	AAH47741 Human hMI
34	1616	61.4	1862	4	AAI59232	AAI59232 Human pol
35	1616	61.4	2380	13	ADQ86789	Adq86789 Human tum
36	1616	61.4	2380	13	ADQ85720	Adq85720 Human tum
37	1616	61.4	2380	13	ACN40520	ACN40520 Tumour-as
38	1612.5	61.3	1605	8	ABT19345	ABt19345 Aspergill
39	1608.5	61.1	1605	8	AAH21165	AAh21165 Aspergill
40	1602	60.9	1833	4	AAH47740	AAh47740 Murine m
41	1602	60.9	2280	1	ABU17361	ABu17361 Drosophi
42	1565.5	59.5	1980	12	ADQ00499	Adq00499 Novel hum
43	1565.5	59.5	1980	12	ADN98930	Adn98930 Novel hum
44	1553	59.0	1991	12	ADG73735	Adg73735 Aspergill
45	1546	58.8	1704	8	ABT20567	ABt20567 Aspergill

ALIGNMENTS

RESULT 1	
ID	ADQ14500
ADQ14500	standard; cDNA; 1533 BP.
XX	
AC	ADQ14500;
XX	
DT	23-SEP-2004 (first entry)
XX	
DE	Mutant soybean myo-inositol 1-phosphate synthase cDNA #3.
XX	
KW	Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW	myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW	inorganic phosphate; mutant.
XX	
OS	Glycine max.
OS	Synthetic.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1533
FT	/*tag= a
FT	/product= "Mutant soybean myo-inositol 1-phosphate
FT	synthase #3"
XX	
FN	US2004128713-A1.
XX	
PD	01-JUL-2004.
XX	
PF	21-NOV-2003; 2003US-00718952.
XX	

PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.

PA	(HITZ/)	HITZ W D.
PA	(SEBA/)	SEBASTIAN S A.
PA	(GRAC/)	GRACE D J.
PA	(STRE/)	STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit IG;

DR WPI; 2004-533135/51.
DR P-PSDB; ADQ14501.

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

PS Claim 10; SEQ ID NO 11; 48bp; English..

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences), where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	3.9e-255
Score:	2631.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	12
Gaps:	0
Matches:	15333
Conservative:	0
Mismatches:	0
Indels:	0

US-10-718-952-12 (1-510) x ADQ14500 (1-1533)

QY		MetPheHisGluAsnPhenylserValGluSerProAsnValysrTrpThrGluIle	20
Db	1	AATGTATCATCAGAAATTATAAGTAGAGAGTCTTAATTTGAAGATACACCGAATCGAATT	60
QY	21	GlnSerValTyrAsnTyrGluThrThrgluLeuValHisGluAsnArgasnGlyThrTyr	40
Db	61	CAGTCGGTGACACTACAGAAAACCACCGAACTTGTTCCAGGAACAGGAATGGCACTTAT	120
QY	41	GlnTrpLleValIleProIlysserValasntyrGlnInpelysrThraantHisValPro	60
Db	121	CAGTGAATTGTCAAAACCCAAATCCGCACAATCAATTTAAACCAACACCACTGTTCCA	180
QY	61	LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnaenglyserThrLeuThrGlyGly	80
Db	181	AAATTGGGGGTGATCGTTGGGGTTGGGGTGGAAACAACGGCTCACCTCACCGGTGT	240
QY	81	ValIleAlaAsnArgGluAspLisestTrpaIatrrLysAspLyfIleGlnGlnAlaAsn	100
Db	241	GTATATGCTAACAGAGAAGACATTTCTATCGGCTAACAAAGACAAGATTCAACAAGCCAAT	300

QY	101	TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu	120
DB	301	TACTTTGGCTCCCTCCACCAGCCTCGACTATTCAGATTGGATCTTCCAGGAGAGAA	360
QY	121	IleTyrAlaProPheIlySerLeuLeuProMetValAsnProAspAlaIleValPheGly	140
DB	361	ATCTATGCCCATTCAGAGAGTCTGCTTCCATGGTTATCTTGACGACATTGCTTTGGG	420
QY	141	GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIlyValPheAsp	160
DB	421	GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTTTGGAC	480
QY	161	IleAspLeuGlnIlySerGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle	180
DB	481	ATCATATTGGCAGAGCAGTTGAGGCTTGAATGAAATCCATGGTTCCATCTCCCGGATC	540
QY	181	TyrAspProAspPheIleAlaAsnGlnGluArgAlaAsnAsnValIleIlySerGly	200
DB	541	TACACCCGATTTTCATTGCTGCCACCAAGAGAGGCTGCCAACAACGTGATTAAAGGC	600
QY	201	ThrIlySerGlnGluGlnValGlnGlnIleIleIlyAspIleIlyValaPheIlySerGluAlaThr	220
DB	601	ACAAAGCAAGGCAAGTTCCAGCAAAATCATCAAAAGACATCAAGGGGTTTAAAGAACCC	660
QY	221	IlyValaIlyAspIlyValaIlyValleuTyrThrAlaAsnThrGluArgTyrSerAsnLeuVal	240
DB	661	AAATGTGACAAAGGTGGTGTCTGTGGACGTGCCAACAGAGAGGTATACCAATTGGTGT	720
QY	241	ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaIlyAlaAspArgAsnGluAlaGlu	260
DB	721	GTAGGCGCTTAATGACACCATGAGAAATCTTGGCTGCTGTGACAGAAATGAGCGTAG	780
QY	261	IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn	280
DB	781	ATTCTCTCTTCCACTTGATGCGCATGCGCTGTGTGTAAGAAATGTCTTTCATTAAAT	840
QY	281	GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr	300
DB	841	GGAAGCCCTCAGAACACTTTTGTATCCAGGGCTGATGATCTTCCATCCGGAGAAACACT	900
QY	301	LeuIleGlyIlyAspAspPheIlySerSerGlyGlnThrIlyMetIlySerValIleValAsp	320
DB	901	TTGATGTGGAGATGACTTCAGAGGTGTGCAGCCAAATGAAATGTGTGTGGTTGAT	960
QY	321	PheLeuValGlyAlaGlyIleIlyProThrSerIleValSerTyrAsnIleGlyAsn	340
DB	961	TTTTTTGGGGGGCTGGTATCAAGCCAAACTCTATATGTTATGAAACCATCTGGGAAAC	1020
QY	341	AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerIlyGlyIleSerIlySer	360
DB	1021	AATGATGGTATGAATCTCTCGGCTCCCAAAACCTTCGCTCCAAAGAAATCTCCAAAGAC	1080
QY	361	AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyIlyIlyAspPro	380
DB	1081	AACGTGTGTGCGATATGTCACAGCAAGATGCCATCTCTATGAGCTGTGTGAACATCCC	1140
QY	381	AspIleValValIleIlyTyrValProTyrValGlyAspSerIlyAspAlaMetAsp	400
DB	1141	GACCATGTGTGTATTAAGTATGTGCTTAAAGGGGATAGCAAGAGCATGATGAT	1200
QY	401	GluTyrThrSerGlnIlePheMetGlyIlyIlyAsnThrIleValLeuIleAsnThrCys	420
DB	1201	GAGTACACTTCAGAGATATTCATGCGGTGGAAGAACCACCTGTTTGGCAACACACTGT	1260
QY	421	GluAspSerLeuLeuAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer	440
DB	1261	GAGAGATTCCTTTAAGCTGCTCTATTAATCTGGACTGTGCTTCTGTGCTGAGCTGAGC	1320
QY	441	ThrTyrGlnGlnPheIlyValaGluAsnGluGlyIlyAspPheIlySerPheIlyAspValAla	460
DB	1321	ACTGAAATCCAGTTTAAAGCTGAAAATGAGGAGAAATTCACATCATTCACCCAGTTGGCT	1380
QY	461	ThrIleLeuSerTyrLeuThrIlyValaProLeuValProProGlyThrProValValAsn	480

DB 1361 ACCATTCTAGCTATCTAGCCAAAGCTCCTCTGTTCCACCGGTAACCACTGGTGAT 1440
QY 461 AAlaSerLeYgAlaArgAlaMetLeuGluAuaIlMeLeArgAlaCyValGlyLeuAla 500
DB 1441 GCATTGTCAAAACAGCGTGCATGCTGGAAACATATAGAGGGCTTGTGTGATTGGCC 1500
QY 501 ProGluAuaAuaMetLileuGluTyrLys 510
DB 1501 CCAGAGAAATACATGATTTCTCAGTACAAAG 1530
RESULT 2
ADS82001 ID ADS82001 standard; cDNA; 1533 BP.
XX ADS82001;
XX 18-NOV-2004 (first entry)
XX Soybean myo-inositol 1-phosphate synthase cDNA mutant #2.
DE Soybean myo-inositol 1-phosphate synthase; ss; gene; phylic acid;
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phylic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
XX Glycine max; line 29010CP01.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 1..1533
XX FT /*tag= a
XX FT /product= "myo-inositol 1-phosphate synthase"
XX FT /replce(260,G)
XX FT /*tag= b
XX mutation
XX US2003074685-A1.
XX 17-APR-2003.
XX 11-MAR-2002; 2002US-00025003.
XX 08-APR-1997; 97US-00835751.
XX PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
PI
XX MPI: 2004-639957/62.
XX P-PSDB; ADS82002.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX phylic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 8; SEQ ID NO 11; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
XX soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX phosphate synthase having decreasing capacity for the synthesis for myo-
XX inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX complement, subfragment or the complement of the subfragment, operably
XX linked to suitable regulatory sequences, where expression of the chimeric
XX gene results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX comprising the chimeric gene (with a heritable phenotype of a seed phylic
XX acid content of less than 17 micromol/g, a seed content of raffinose plus
XX stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX greater than 200 micromol/g, provided that the plant is not LR33), seeds
XX from the plant, making a soybean plant with the heritable phenotype
XX (comprising crossing LR33 or the plant comprising the chimeric gene with
XX an elite soybean plant and selecting a progeny plant of the cross of

CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phylic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a mutant myo-
CC inositol 1-phosphate synthase.
XX
XX Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 3.9e-255 Length: 1533
XX Score: 2631.00 Matches: 510
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 13 Gaps: 0
XX
XX US-10-718-952-12 (1-510) x ADS82001 (1-1533)
QY 1 MetPheIlleGluAuaAuaPheLysValGlySerProAuaValLysTyrThrGluThrGluIle 20
DB 1 ATGTCATCGAAGATTTTAAGTAGAGAGTCTTAATGGAAGTACACGAGCTAGATT 60
QY 21 GlnSerValTyrAuaTyrGluThrThrGluLeuValHisGluAuaArgAnglyThrTyr 40
DB 61 CAGTCGGTACAACTACGAAACCAACCACTGTTCACGAAACGAAATGGCACCTAT 120
QY 41 GlnTrpIleValLysProLysSerValAuaTyrGlnPheLysThrAsnThrHisValPro 60
DB 121 CAGTCGATTCGAAACCAACCAATCCCTCAACTACCAATTTAAACCAACCACTTTCCA 180
QY 61 LysLeuGlyValMetLeuValGlyTyrGlyLysAuaAnglySerThrLeuThrGlyGly 80
DB 181 AAATTGGGGGATGATCTTGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGG 240
QY 81 ValIleAlaAuaArgGluAspIleSerTrpAlaThrLysAuaPheLysIleGlnIleAlaAua 100
DB 241 GTTATTGCTCAACAGAGACATTTCAATGAGGCTTCAAAAGCAATTCACAAACCAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGlu 120
DB 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCAGATTGGATCTTCCAGGAGAGAGA 360
QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAuaProAspAspIleValPheGly 140
DB 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAAATGTTAATCCGAGCATTTGTTGGG 420
QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGAGATATCAGCAACATGAACTGCTGATGCAATGCGGCAAGGCAAGGTTGTTAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
DB 481 ATCGATTTGCGAAGACAGGTGAGGCTTACATGATCAATGATCAATGATCTCCCGGAAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 541 TACGACCCGGATTTTATTGCTGCTGCAACCAAGGAGGCGTGCACCAAGTATTAAAGGCG 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleLysAspIleLysValAuaPheLysGluAlaThr 220
DB 601 ACAAAGCAAGAGCAAGTTCACCAATATCTCAAAAGCAACCAAGCGCTTTAAGAAACCCAC 660
QY 221 LysValAuaPheLysValValIleLeuTrpThrAlaSerThrGluArgTyrSerAsnLeuVal 240

```
Db 661 AAAGTGACAAAGTGTTGCTGCTGAGCTGCCAACACAGAGAGGTATAGCAATTGGTT 720
Qy 241 ValGlyLeuAsnAspThrMetGluLeuLeuAlaValAspArgAsnGlu 260
Db 721 GTAGGCTTAAATGACCACTGAGAAATCTCTGGCTGCTGAGACAGAAATGAGCTGAG 780
Qy 261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 781 ATTCTCCTCCATCCATGATGATGATGCTGTGTGATGAGAAATGTTCTTCAATTAAT 840
Qy 281 GlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 841 GGAACCCCTCAGAACACTTTGTACAGGCTGATGTGATCTTGCCATGCGCAGGAACACT 900
Qy 301 LeuIleGlyValAspAspPheLeuSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 901 TTGATTGGTGGAGAGACTTCACAGAGTGTCTAGACCAAAATGAAATCTGTGTGGTTGAT 960
Qy 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db 961 TTTCTTGTTGGGGCGTGGATACAGCAACATCTATAGTTAGTTACACACATCTGGGAAC 1020
Qy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db 1021 AATGATGGTATGAATCTCTCGGCTCCACAAACCTTCGCTCCAGGAAATCTCCAAAGAC 1080
Qy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyValHisPro 380
Db 1081 AACGTTGTTGACGATATGCTCAACAGCANTGCCATCTCTATGAGCCTGGTGAACATCCC 1140
Qy 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db 1141 GACCATGTTGTTGTTATTAAATGATGCTTACGTAGGGGATAGCAAGAGACCATGAT 1200
Qy 401 GluTyrThrSerGluIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys 420
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAAGAACACCATGTTTGGACAAACATGT 1260
Qy 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1261 GAGGATTCCTTTAGCTGCTCTTATTCCTTGACCTTGCTCTTCTTGCTAGCTGAGC 1320
Qy 441 ThrArgIleGlnPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
Db 1321 ACTAGATTCACAGTTTAAAGCTGAAATAGGAAATTCACATTCACACCATGTTGCT 1380
Qy 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db 1381 ACCATTCTCAGCTATCTGACCAAGGCTCTCTGCTGCTCCACCGGATACACCACTGGTGAAT 1440
Qy 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1441 GCATTGTCAAGACGAGTGCATAGCTGGAAACATATGAGGGCTTGTGTGATTGGCC 1500
Qy 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
Db 1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAG 1530
RESULT 3
ADQ14504
ID ADQ14504 standard; cDNA; 1533 BP.
XX AC
XX ADQ14504;
XX
XX 23-SEP-2004 (first entry)
XX
XX wild type soybean myo-inositol 1-phosphate synthase cDNA #2.
XX
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate.
XX
XX Glycine max.
OS
```

```
XX Key Location/Qualifiers
FH 1..1533
FT CDS
FT
FT
FT
FT US2004128713-A1.
PN
XX
XX 01-JUL-2004.
PD
XX
XX 21-NOV-2003; 2003US-00718952.
PF
XX
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98MO-US006822.
PR 26-APR-1999; 99US-0029315.
PR 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
PI
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX WPI; 2004-533135/51.
DR
XX P-PSDB; ADQ14505.
PT
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
PS
PS Claim 4; SEQ ID NO 15; 48pp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents cDNA encoding a wild type
XX soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
XX SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1,98e-254 Length: 1533
XX Score: 2624.00 Matches: 509
XX Percent Similarity: 99.80% Conservative: 0
XX Best Local Similarity: 99.80% Mismatches: 1
XX Query Match: 99.73% Indels: 0
XX DB: Gaps: 0
XX
XX US-10-718-952-12 (1-510) x ADQ14504 (1-1533)
Qy 1 MetPheIleGluAsnPheLysValGluSerProAsnValLysTyrThrGluThrGluIle 20
Db 1 ATGTCATCGAGAAATTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAAGCTGAGATT 60
Qy 21 GlnSerValIlyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 61 CAGTCCGTGTAACAATAAGCAACCAAGCACTTGTTCAGAGAACAGGAATGGCACTTAT 120
```


QY 41 GlnTrpIleValIleAspProIleSerValIleAsnTyrGlnPheIleThrAsnThrHisValPro 60
 Db 121 GAGTGGATGTCAGAACCCAAATCCGTCACTACCAATTTAAACCAACCCAGTTTCCA 180
 QY 61 LysLeuGlyValMetLeuValGlyTrpGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
 Db 181 AATTGGGGGTGATGTTGTGGGTGGGGAACAACGGCTTACCTCCACGGTGTGT 240
 QY 81 ValIleIleAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnGlnIleAsn 100
 Db 241 GTTATTGCTAACAGAGGGCATTTTCATGGGCTAACAGGACAAAGATTCAACACGCAAT 300
 QY 101 TyrPheGlySerLeuThrGlnIleAsnAlaIleArgValGlySerPheGlnGlyGln 120
 Db 301 TACTTGGCTCCCTCCCAAGCCCTCAGCTATTGAGTTGAGTCTTCCAGGGAGAGGA 360
 QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
 Db 361 ACTATGCCCCCATTTCAAGAGTCTGCTTCCATGGTTATCTTGAACATGTGTGGG 420
 QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
 Db 421 GGATGGATATCAGCAACATGAACCTGGCTGATGCTGCGCCAGGCAAGGTGTTGAC 480
 QY 161 IleAspLeuGlnLysGlnLeuArgProTyrTrpMetGluSerMetValProLeuProGlyIle 180
 Db 481 ATCGATTTCAGAGGAGGTGAGGCTTACATGAAATCCATGGTTCCACTCCCGGAAATC 540
 QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluGluArgAlaAsnAsnValIleLysGly 200
 Db 541 TACGACCGGGATTTCATTGCTGTCACCAAGAGGACGGTCCCAACACGTGATTAAGGGC 600
 QY 201 ThrLysGlnGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
 Db 601 ACAAAAGCAAGACCAAGTTGACGAATCATCAAGACATCAAGGCGTTTAAAGAGGACACC 660
 QY 221 LysValAspLysValValIleLeuTrpTrpAlaAsnThrGluArgTyrSerAsnLeuVal 240
 Db 661 AAGGTGGACAGGGTGTGTCTGCTGAGCTGCGCAACAGAGGATGTCATTTGGTT 720
 QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaAlaValAspArgAsnGluAlaGlu 260
 Db 721 GTAGGCTTAAATGACCACTGAGATCTCTTGCGCTGTGACAGAAATGAGGCTGAG 780
 QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
 Db 781 ATTTCTCCTTCCACCTTGTATGCAATTCCTGTGTGATGGAATAATGTTCTTTCATTAT 840
 QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
 Db 841 GGAAGCCCTCAAGAACCTTTGTATCCAGGGCTGATTTGATCTTCCAGGAGACACT 900
 QY 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
 Db 901 TTGATTGGTGGAGATCACTTCAAGAGTGTCAAGCAAAATGAAATCTGTGTGGTAT 960
 QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisIleuGlyAsn 340
 Db 961 TTTCTTGGGGGCTGTGTATCAAGCCACATCATATGTTAGTTACAAACCATCTGGGAAAC 1020
 QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
 Db 1021 AATGATGGTAAATCTCTGGCTCCAAACCTTCCGCTCCAAAGGAAATCTCCCAAGAC 1080
 QY 361 AsnValIleAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
 Db 1081 AACGTTGTGAGATATGCTCAACAGCATTCCTCTTATGAGCTGTGAAATCC 1140
 QY 381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
 Db 1141 GACCATGTTGTATTAATGATATGCTTACGTAGGGGATAGCAAGAGACCATGAT 1200

QY 401 GlnTyrThrSerGluIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys 420
 Db 1201 GAGTACACTTCAGAGATATTCAATGGGTGAAAGACACCATTTGTTGCAACACATGT 1260
 QY 421 GluAspSerLeuLeuAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
 Db 1261 GAGGATTCCTTTTGGTCTCTCTATTAATCTTGAGACTTGCTCTTCTTGAGCTGAGC 1320
 QY 441 ThrArgIleGlnPheLysAlaGluAsnGlyLysPheHisSerPheHisProValAla 460
 Db 1321 ACTAGATCCAGTTTAAACCTTAAATGAGGAAATTCACATTCACCCAGTTGCT 1380
 QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyTyrProValValAsn 480
 Db 1381 ACCATTCTAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGATACACAGTGTGAAT 1440
 QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
 Db 1441 GCATTGTCAAACAGGTCATGCTGGAATAATGAGGCTTGTGTGATTTGGATGGCC 1500
 QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
 Db 1501 CCGAGAAATTAACATGATCTCGAGTACAG 1530
 RESULT 4
 ADQ14502
 ID ADQ14502 standard; cDNA; 1533 BP.
 AC ADQ14502;
 XX
 DT 23-SEP-2004 (first entry)
 XX
 DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #4.
 XX
 KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
 KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
 XX inorganic phosphate; mutant.
 OS Glycine max.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT 1..1533
 FT CDS /tag= a
 FT /product= "Mutant soybean myo-inositol 1-phosphate
 XX synthase #4"
 PN US2004128713-A1.
 XX
 PD 01-JUL-2004.
 PD
 PF 21-NOV-2003; 2003US-00718952.
 PF
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 PR 26-APR-1999; 99US-00293315.
 PR 11-MAR-2002; 2002US-00025003.
 PR
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PA (GRAC/) GRACE D J.
 PA (STRE/) STREIT L G.
 PA
 PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
 PI
 XX WPI: 2004-533135/51.
 DR P-PSDB; ADQ14503.
 DR
 XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytic acid and increased sucrose, leading to valuable and useful soybean
 PT products.
 XX

PS Example 8; SEQ ID NO 13; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,98e-254	Length:	1533
Score:	2624.00	Matches:	509
Percent Similarity:	99.80%	Conservative:	0
Best Local Similarity:	99.80%	Mismatches:	1
Query Match:	12	Indels:	0
DB:		Gaps:	0

US-10-718-952-12 (1-510) x ADQ14502 (1-1533)

```
QY 1 MetPheIIeGluAsnPhelyValGIuserProaenValIySerThrGIunhrgIuile 20
DB 1 AAGTTCATCGAGAATTTTAAGGAGAGAGTCCTAAGTGAAGACCCAGACTGAGATT 60
QY 21 GInSeRValTYrAsnTYrGIunhrgIuValhIeGIuAsnArGsnGIYhTYr 40
DB 61 CAGTCCGCTACACACGAAACCCGAACTTTGTCACGAGAACGAAATGGACCTTAT 120
QY 41 GInTPRIleValIyProIySeRValAenTYrGInPhelySerThrsnThiSValPro 60
DB 121 CAGTGGATGTCAAACCCAAATCCGTCACATCTTAAACCAACCCATGTTCCA 180
QY 61 LysIeuGIyValMetIeuValGIYrPGIyGIyAsnSngIySeRThrsnThiSValPro 80
DB 181 AAATTGGGGGTATGCTTGTGGGTGGGTGGAAACACGCGCTTACCCGCTGCT 240
QY 81 ValIleAlaAsnArGIuAspIleSeRTPrAlaThrIyAspIySIIeGIInGIuAlaAsn 100
DB 241 GTTATTGCTACACGAGGCGCATTTCAATGGCTACAAAGACAAAGTTCAACAGCAT 300
QY 101 TYrPheGIySeRleuThrgInIaSeRAlaIleArgValGIySeRPhelngIyGIuIn 120
DB 301 TACTTTGGCTCCCTCCACCAAGCCTCAGCTATTCCGAGTGGATCTTCCAGGAGGANA 360
QY 121 ILeTYrAlaProPhelySeRleuPROMetValAsnProAspIleValPheGIy 140
DB 361 ACTTATGCCCATCTCAAGAGTGTGCTTCCAAATGGTTAATCTGCAGCAACATGGTTGGG 420
QY 141 GIYTPAspIleSeRAsnMetAsnIeuAlaAspAlaMetAlaArgAlaIyValPheAsp 160
DB 421 GGATGGGATATCGACAACTGAACCTGCTGATGCATGCCAGGCAAAAGGTTTGCAC 480
QY 161 ILeAspLeuGIuIySngIuIeuArgProTYrMetGIuSeRMetValProIeuProGIyIle 180
DB 481 ATCGATTTCAGAAAGAGCTTGAAGCTTACATGGAATTCATGGTTCACCTCCCGGAATC 540
QY 181 TYrAspProAspPheIleAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 200
DB 541 TACGACCCGGAATTCATGCTGTGCCAACCAAGAGAGAGCGTCCAAACGATGATTAAGGCGC 600
```

```
QY 201 ThrIySngIuGIuValGIInGIuIleIleYsAspIleIySValaPheIyGluAlaThr 220
DB 601 ACAAAGCAAGAGCAAGTTCAAGAAATCATCAAGACATCAAGGCGTTTAAAGAAAGCCACC 660
QY 221 LysValaAspIySValaValaIeuThrPhrAlaSnThrgIuArgTYrSeRAsnIeuVal 240
DB 661 AAAGGCAAGAGGTGTCTCTGGAGCTCCCAACAGAGAGGTATAGCAATTTGGT 720
QY 241 ValGIyIeuAsnAspThMeGIuAsnIeuIuAlaIaValaAspArGsnGIuIaGIu 260
DB 721 GTAGGCTTAATGACACCATGAGAAATCTTGGCTGTGTGACAGAAATGAGGCTAG 780
QY 261 ILeSeRProSeRThrIeuTYrAlaIleAcYsValMetGIuAsnValProPheIleAsn 280
DB 781 ATTCTCTTCACCTGTATGACATTCCTGTGTGATGAGAAATGTCTCTTCAATTAT 840
QY 281 GIYSeRProGIuAsnThrPheValProGIyIeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAACCTTCAGAAACCTTTGTACCAAGGCTGATGATCTTGGCCATGCGAGAAACACT 900
QY 301 LeuIleGIyIyAspAspPheIySeRGIyInThrIyMetIySeRValIeuValAsp 320
DB 901 TTGATGGTGGAGATGACTTCAAGAGTGTGCAGACCAAAATGAATCTGTGTGGTAT 960
QY 321 PheIeuValGIyAlaGIyIleYsProThrSeRIleValSeRTYrAsnHIsIeuGIyAsn 340
DB 961 TTTCTTGGGGGCTGTATCAAGCAACATCTATAGTTTACAAACATCTGGGANAAC 1020
QY 341 AsnAspGIyMetAsnIeuSeRAlaProGIunhrgPheArgSeRGIySngIySeRtySeR 360
DB 1021 AATGATGGTATGAATCTTCGCTCCACAAACCTTCGCTCCAGAAATCTCCAAAGAGC 1080
QY 361 AsnValIaAspAspMetValAsnSeRAsnAlaIleIeuTYrGIuProGIyGIuHIsPro 380
DB 1081 AACGTTGTGACGATATGTCACACGACATGCCATCTCTATGAGCCTGGTGAACATCC 1140
QY 381 AspHIsValaIleYsTYrValProTYrValGIyAspSeRtySValaMetAsp 400
DB 1141 GACCATGTTGTGTATTAAGTATATGTGCTTACGTAAGGAGTATGACAAAGACCATGAT 1200
QY 401 GIuTYrThrSeRGIuIlePheMeGIyGIyIySAsnThrIleValIeHIsAsnThrgYs 420
DB 1201 GAGTACACTTCAGAAATTCATAGGGTGAAAGAACACCATGTTTTCACAAACATATG 1260
QY 421 GIuAspSeRleuIeAlaProIleIleIeuAspIeuValIeuIeuAlaGIuIeuSeR 440
DB 1261 GAGATTCCTTTTATGCTGCTCTATTTATCTTGACCTTGCTCTTGCTGAGCTGAGC 1320
QY 441 ThrArgIleGIuPheIySValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 460
DB 1321 ACTAGAAATCCAGTTTAAAGCTGAATAATGAGGAAATTCACATCTCCACCCAGTTGCT 1380
QY 461 ThrIleIeuSeRTYrIeuThrIySValaProIeuValProProGIyThrProValaIaSn 480
DB 1381 ACCATTCTCAGTATCTTACCAAGGCTCTCTGTGTCCACCGGATACACCATGTGTAAAT 1440
QY 481 AlaIeuSeRTySngIuArgAlaMetIeuGIuSnIleMetArgAlaCySValaGIyIeuAla 500
DB 1441 GCATTGTCAAAGCAGCGTGCATATGCTGAAAAACATTAAGAGGCTGTGTGTGATTTGCC 1500
QY 501 ProGIuAsnAsnMetIleIeuGIuTYrIyS 510
DB 1501 CCAGAGAAATTAACATGATTTCTCGAATGAAG 1530
RESULT 5
ADSS82005 standard; cDNA, 1533 BP.
ADSS82005;
ADSS82005;
18-NOV-2004 (first entry)
Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 2.
```

XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phylic acid;
 KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
 XX Glycine max; cultivar Wye.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1533
 FT /tag= a
 FT /product= "myo-inositol 1-phosphate synthase"
 XX
 PN US2003074685-A1.
 XX
 PD 17-APR-2003.
 XX
 PF 11-MAR-2002; 2002US-00025003.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PI Hitz WD, Sebastian SA;
 XX
 DR MPI: 2004-63957/62.
 DR P-PSDB: ADS82006.
 XX
 PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
 PT phylic acid and inorganic phosphate content of soybean seeds.
 XX
 PS Claim 2: SEQ ID NO 15; 34pp; English.
 CC The invention relates to an isolated nucleic acid fragment encoding a
 CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 CC phosphate synthase having decreasing capacity for the synthesis for myo-
 CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
 CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 CC complement, subfragment or the complement of the subfragment, operably
 CC linked to suitable regulatory sequences, where expression of the chimeric
 CC gene results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 CC comprising the chimeric gene (with a heritable phenotype of a seed phylic
 CC acid content of less than 17 micromol/g, a seed content of raffinose plus
 CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
 CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype
 CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homozygous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with a
 CC heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phylic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence encodes a wild-type myo-
 CC inositol 1-phosphate synthase.
 CC
 XX
 SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.98e-254 Length: 1533
 Score: 2624.00 Matches: 509
 Percent Similarity: 99.80% Conservative: 0
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 99.73% Indels: 0

DB: 13 Gaps: 0
 US-10-718-952-12 (1-510) x ADS82005 (1-1533)
 QY 1 MetPheileGluAsnPhelyeValGlySerProAsnValTyrThrGluThrGluile 20
 DB 1 ATGTTTCATCGAATAATTTTAAGGTAGAGAGTCTTAATGTGAATCAACGAGCTGAGATT 60
 QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHleGluAsnArgAsnGlyThrTyr 40
 DB 61 CAGTCCGTGTACAACTACGAAACCCGAACTTGTTCAGAAACAGGAATGGACCTAT 120
 QY 41 GlnTrpIleValLysProLysSerValAsnTyrGlnPheLysThrAsnThrHisValPro 60
 DB 121 CAGTGATGTTCAAACCCAAATCCGTCAACTACAAATTTAAACCAACCCCATGTTCGA 180
 QY 61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
 DB 181 AAATTGGGGGTGATCTTGTGGGTGGGTGGAACCAACGCTTACCTCACCGGTGT 240
 QY 81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
 DB 241 GTTATTGCTTAACAGAGAGGCAATTCATGGGCTTCAAAAGACAAAGATTCAACAGCCAA 300
 QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
 DB 301 TACTTGGCTCCCTTACCCAAAGCTCAGCTATTCAGATTGATCTTCAGAGAGAGNA 360
 QY 121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
 DB 361 ATCTATGCCCTTCAAGAGTCTGCTTCAATGTGTTAATCTCGAGACATTTGTGGG 420
 QY 141 GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
 DB 421 GGATGGATATACGAACATGAACCTGGCTGATGCCATGGCCAGAGGCAAAAGGTGTTGAC 480
 QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGlySerMetValProLeuProGlyLys 180
 DB 481 ATCGATTTGCAAGAACCACTGTGAGCTTACATGAAATCATGATGTTCCACTCCCGAATC 540
 QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluValArgAlaAsnValIleLysGly 200
 DB 541 TAGACCCGGATTTTCATTTGCTGCCCAACAGAGAGCTGCCAACAAGCTATTAAAGGCC 600
 QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
 DB 601 ACAAAGCAAGAGCAAGTTCAGCAATTCATCAAGACATCAAGAGCGTTTAAAGAACCCAC 660
 QY 221 LysValaAspLysValValValLeuThrThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
 DB 661 AAAGTGCACAAAGTGTGTCTCTGTGACTGCCAACACAGAGAGGTATAGCAATTTGGTT 720
 QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValaAspArgAsnGluAlaGlu 260
 DB 721 GTAGGCTTTAATGACACATGAGATCTTTGGCTGCTGTGACAGAAATGAGCTTAG 780
 QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
 DB 781 ATTTCTCTTCACCTGTATAGCCATTCCTGTGTGATGGAATAAGTTCTTTTCAATTAT 840
 QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
 DB 841 GGAAACCTCCAGAACACTTTGTGTACAGAGGCTGATGTCTTGGCATGCGCAGGAACACT 900
 QY 301 LeuIleGlyLysAspPheLysSerGlyGlnThrLysMetLysSerValLeuValaAsp 320
 DB 901 TTGATGTGTGAGATGACTTCAAGAGTGTGACAGCAAAATGMAATCTGTGTGTTAT 960
 QY 321 PheLeuValGlyAlaGlyLysPheProThrSerIleValSerTyrAsnHleGluLys 340
 DB 961 TTTCTTGGGGGGCTGATATCAAGCAACATCTATAGTTAGTTACCAACCATCTGGGAAC 1020
 QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360

```

DB      1021 AATGATGGATGATGATCTCGGCTCCACAAACCTTCGCTCCAGAAATCTCCAGAGC 1080
QY      361 AenVal1ValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB      1081 AACGTTGTTGAGCATRATGCTCAACAGCATGCCATCTCTTAAGAGCTGTGTAACATCCC 1140
QY      381 AspHisValValIleLeuTyrValProTyrValGlyAspSerIysArgAlaMetAsp 400
DB      1141 GACCATGTTGTTGATTATTAATGATGCTTACGTAGGGGATAGCAAGAGACCATGGAT 1200
QY      401 GluTyrThrSerGluIlePheMetGlyGlyAsnThrIleValLeuHisAsnThrCys 420
DB      1201 GAGTACACTTCAGAGATATTCATGGGTGAAACACACCATGTTTGCACAACACATGT 1260
QY      421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValIleLeuAlaGluLeuSer 440
DB      1261 GAGGATTCCTTTTACCTCTCTCTATTATCTTGACCTTGCTCTTCTTCTAGCTGACC 1320
QY      441 ThrArgIleGlnPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
DB      1321 ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCACACTCATTCACCCAGTTGCT 1380
QY      461 ThrIleLeuSerTyrIleuThrIleLysAlaProLeuValProProGlyThrProValValAsn 480
DB      1381 ACCATTCTCAGCTATCTGACCAAGGCTCCTCTGTTCCACCGGTACACACAGTGTGAT 1440
QY      481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB      1441 GCAATTGTCAAAACAGCGTCCAAATGCTCGAAGAAACATATAGGGCTTGTGTGGATTGGCC 1500
QY      501 ProGluAsnAsnMetIleLeuGluTyrIlys 510
DB      1501 CCAGAGAAATACATGATTTCTCGAGTACAAG 1530

RESULT 6
ADS82003
ID      ADS82003 standard; cDNA; 1533 BP.
AC      ADS82003;
XX      18-NOV-2004 (first entry)
DT      Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 4.
DE      Soybean myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
KW      raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX      Glycine max; line 29018UP03.
OS      Glycine max; line 29018UP03.
FH      Key Location/Qualifiers
FT      CDS 1..1533
FT      /product= "myo-inositol 1-phosphate synthase"
FT      /tag= a
FT      /product= "myo-inositol 1-phosphate synthase"
XX      US2003074685-A1.
XX      17-APR-2003.
XX      11-MAR-2002; 2002US-00025003.
XX      08-APR-1997; 97US-00835751.
XX      07-APR-1998; 98WO-US006822.
XX      (HITZ/) HITZ W D.
XX      (SEBA/) SEBASTIAN S A.
XX      Hitz WD, Sebastian SA;
XX      PI
XX      DR WPI: 2004-639957/62.
XX      P-PSDB; ADS82004.
XX      Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-

```

```

PT      phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT      phytic acid and inorganic phosphate content of soybean seeds.
XX      Example 8; SEQ ID NO 13; 34pp; English.
CC      The invention relates to an isolated nucleic acid fragment encoding a
CC      soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC      phosphate synthase having decreasing capacity for the synthesis for myo-
CC      inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC      nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC      complement, subfragment or the complement of the subfragment, operably
CC      linked to suitable regulatory sequences, where expression of the chimeric
CC      gene results in a decrease in expression of an endogenous or native gene
CC      encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC      comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC      acid content of less than 17 micromol/g, a seed content of raffinose plus
CC      stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC      greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC      from the plant, making a soybean plant with the heritable phenotype
CC      (comprising crossing LR33 or the plant comprising the chimeric gene with
CC      an elite soybean plant and selecting a progeny plant of the cross of
CC      crossing step that has a heritable phenotype as mentioned above), seeds
CC      of soybean plant made by the above method, a soy protein product derived
CC      from seeds of a soybean plant (homozygous for one or more gene encoding a
CC      mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC      the synthesis of myo-inositol 1-phosphate), where the gene confers a
CC      heritable phenotype as mentioned above), and making or producing a
CC      soybean protein product derived from seeds of a soybean plant with
CC      heritable phenotype as mentioned above. The nucleic acid is useful for
CC      altering raffinose saccharide, sucrose, phytic acid and inorganic
CC      phosphate content of soybean seeds thus leading to valuable and useful
CC      soybean products, since the presence of high concentration of raffinose
CC      oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC      when consumed by humans. The present sequence encodes a wild-type myo-
CC      inositol 1-phosphate synthase.
SQ      Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,98e-254 Length: 1533
Score: 2624.00 Matches: 509
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.73% Indels: 0
DB: 13 Gaps: 0

US-10-718-952-12 (1-510) x ADS82003 (1-1533)
QY      1 MetPheIleGluAsnPheLysValGluSerProAsnValIysTyrThrGluThrGluIle 20
DB      1 ATGTTCAATCGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAAGCTGAGATT 60
QY      21 GlnSerValIysAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB      61 CAGTCCGTTGACAACTACGAAACCCGAACTTGTTACGGAACAGGAATGGACATTAT 120
QY      41 GlnTrpIleValIysProLysSerValAsnTyrGlnPheLysTrpAsnThrHisValPro 60
DB      121 CAGTGGATGTCAAACCAAAATCCGTCAACTACCAATTTAAACCAACACCAAGTTTCA 180
QY      61 LysLeuGlyValMetLeuValGlyTyrGlyGlyAsnAngIysSerThrLeuThrGlyGly 80
DB      181 AAATGGGGGTGATGCTTGTGGGTGGGTGGAACCAACGCGCTTACCTCACCGGTGGT 240
QY      81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnAlaAlaAsn 100
DB      241 GTTATGCTAACAGAGAGGCGCATTTTCAGGGCTTCAAAAGCAAGATTCAACAGCCAT 300
QY      101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB      301 TACTTTGGCTCCCTCACCACCAAGCCTCAGCTATTGCAAGTTGATCTCTCCAGGAGAGAA 360
QY      121 IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspIleValPheGly 140

```

```

Db      |||
361 ACTATGCCCCCATTCAGAGTGTGCTTCCATGTGTTAATCTGTGACGACATTTGTTGGG 420
Qy      |||
141 GYTPAPPII6SerAsnMetAsnLeuA1aAspAlaMetAlaArgAlaIysValPheAsp 160
Db      |||
421 GGATGGGATATCAGCAACATGAACCTGGCTGATGGCATGGCCAGGCGCAAGGTTGTGAC 480
Qy      |||
161 ILeapLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
Db      |||
481 ATCGATTTTCAGAGGAGATTGAGGCTTACATGGAATCCATGGTTCACTCCCGGGAATC 540
Qy      |||
181 TYAspProAspPheIleAlaIaAsnGlnGluArgAlaAsnAsnValIleIysGly 200
Db      |||
541 TACGACCCGGATTTCACTTCTGTGCAACAGAGGAGCGGCCAACAGCATGATTAGGGC 600
Qy      |||
201 ThrLysGlnGluValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
Db      |||
601 ACAAAAGCAAGACCAAGATTCAAGCAAAATCATCAAGATCAAGCGCTTTAAGGAGCCACC 660
Qy      |||
221 LysValAspLysValValIleLeuThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db      |||
661 AAGTGGACAGAGTGTGTCTGTGACTGCCACACAGAGGATAGCAATTTGGTT 720
Qy      |||
241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu 260
Db      |||
721 GTAGGCTTAATGACACCATGAGAAATCTTGGCTGTGACAGAAATGAGGCTGAG 780
Qy      |||
261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db      |||
781 ATTTCTCCCTTCCACCTTGTATGCAATTCCTGTGTATGGAATAATGTTCTTTCATTAA 840
Qy      |||
281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      |||
841 GGAAGCCCTCAAGAACCTTTGTACCAAGGCGCTGATTGATCTTCCATCCGAGGAACACT 900
Qy      |||
301 LeuIleGlyLysAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db      |||
901 TTGATTTGGTGAGATGACCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTTGAT 960
Qy      |||
321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db      |||
961 TTTCTTGTGGGGCTGTGATCAAGCCAACTATATGTTACCAACCATCTGGGGAAC 1020
Qy      |||
341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db      |||
1021 AATGATGATATATCTCTGGCTCCAAACCTTCCTCCAAAGAAATCTCCCAAGAC 1080
Qy      |||
361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlnHisPro 380
Db      |||
1081 AACGTTGTGAGATATGTCACAACAGCATTCCTATGAGCTGTGTAACATCCC 1140
Qy      |||
381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db      |||
1141 GACCATGTTGTATATATGATATGCTTACGTAGGGGATAGCAAGAGACCATGAT 1200
Qy      |||
401 GUYTYrThrSerGlnIlePheMetGlyLysLysAsnThrIleValLeuHisAsnThrCys 420
Db      |||
1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACACATGTTTGCACAACACATGT 1260
Qy      |||
421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db      |||
1261 GAGGATTCCTTTTACCTCTCTATTAATCTTGGATGTGCTCTCTTCTGTCTAGCTGACC 1320
Qy      |||
441 ThrArgIleGlnPheLysAlaGluAsnGlnGlyLysPheHisSerPheHisProValAla 460
Db      |||
1321 ACTAGAAATCAGATTAAAGCTGAATAATGAGGAAATTCACATTCATCCACAGTGGCT 1380
Qy      |||
461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db      |||
1381 ACCATTCTCAGTATCTGACCAAGGCTCTCTGTTCCACCGGATACACACAGTGTGAT 1440
Qy      |||
481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500

```

```

Db      1441 GCATTGTCAAGACAGGTGCAATGCTGGAATAATAGAGGCTTGTGTGATTGCC 1500
Qy      501 ProGluAsnAsnMetIleLeuGlyTyrLys 510
Db      1501 CCAGAAATTAACATGATTTCTCGATGACAAAG 1530
RESULT 7
ADQ14498
ID ADQ14498 standard, cDNA; 1533 BP.
XX
AC ADQ14498;
XX
DT 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #2.
XX
KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate; mutant.
XX
OS Glycine max.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /*tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT /synthase #2"
PN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006882.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
DR WPI: 2004-53135/51.
DR P-PsDB; ADQ14499.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
PS Example 8, SEQ ID NO 9; 48bp; English.
XX
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful

```

CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the inversion.

SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,65e-251	Length:	1533
Score:	2593.00	Matches:	502
Percent Similarity:	99.22%	Conservative:	4
Best Local Similarity:	98.43%	Mismatches:	4
Query Match:	98.56%	Indels:	0
	12	Gaps:	0

US-10-718-952-12 (1-510) x ADQ14498 (1-1533)

```

QY 1 MetPheileGluAsnPhelysValGluSerProAsnValIleYrThrGluThrGluIle 20
DB 1 ATGTCATCGAATAATTTTAAAGTTGAGTGTCTTAATGCAATACCAAGACTGAGATT 60
QY 21 GluSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
DB 61 CAGTCGCTGTACACTACGAAACCAACGAACTTGTTACGAGAACAGGAATGGCACTAT 120
QY 41 GlnTPPleValIlePProIlySerSerValAsnTyrGlnPheIlyThrAsnThiIstValPro 60
DB 121 CAGTGGATTGTCAAAACCCAAATCTGTCAAAATGCAATTTAAACCAACATTCATGTTCT 180
QY 61 LysLeuGluValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
DB 181 AAATTAAAGGGTAATGCTTGTGGGTGGGGTGAAGAACAGCGCTCAACCTCCACCGGTGT 240
QY 81 ValIleIleAsnArgGluAsnIleSerTyrAlaThrLysAsnLysIleGlnGlnAlaAsn 100
DB 241 GTTATTGCTTAACCGAGAGGCAATTCATGCGCTACAAAGACAAAGATTCAACAAACCAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
DB 301 TACTTGGTCCCTCAACCCAAAGCTTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGGAA 360
QY 121 IleTyrAlaProPheIlySerLeuLeuProMetValAsnProAsnAspIleValPheGly 140
DB 361 ATCTATGCCCAATTCAGAGGCTGCTCCAAATGCTTAACCTGACACACTGTGTGGG 420
QY 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaIleValPheAsp 160
DB 421 GGATGGGAATACAGCAACATGAACCTGGCTGATGCATGGCCAGGCAAGGTGTTGAC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
DB 481 ATCGATTGTCAAGAGAGTTGAGGCTTTCATGGAATTCATGCTCCACCTCCGGAATC 540
QY 181 TyrAspProAspPheIleAlaIleAsnGlnGluGluArgAlaAsnAsnValIleLysGly 200
DB 541 TATGACCCCGATTTCATTCCTGCTCCAAACAGAGAGCGGCCAAACACTCATCAAGGCG 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
DB 601 ACAAAAGCAAGACCAAGTTCAACAATCATCAAAAGCATCAAGAGCGCTTAAAGAAACCAAC 660
QY 221 LysValAspLysValValValLeuThrPThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
DB 661 AAAGTGAACAGAGTGTGTACTGTGAGCTGCAACACAGAGAGTACAGTAATTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaIleValAspArgAsnGluAlaGlu 260
DB 721 GTGGGCTTAAATGACCACTGAGAAATCTCTTGCTGCTGCTGTGACAGAAATAGGCTGAG 780
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTTGCTCTTCCACCTGTGATGCATGTGCTGTGATGAAAGAAAGTTCTTCAATTAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300

```

```

DB 841 GGAAACCTCTCAACACCTTTTGTACCGAGGCTGATTGATCTTGGCATCGCAGGAACACT 900
QY 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValIleuValAsp 320
DB 901 TTGATTGGTGAAGATGACTTCAAGAGTGGTCAAGCAAAATGAATCTGTGTGGTTGAT 960
QY 321 PheLeuValGlyAlaGlyIleLysPProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
DB 961 TTCTTGGGGGGCGTGGTATCAACCAACATCTATAGCACTTACCAACATCTGGGAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
DB 1021 AATGATGATGAATCTTTCGGCTCCACAACTTTCGTTCCAGGAATCTCCAAAGGC 1080
QY 361 AsnValValaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
DB 1081 AACGTTGTGATGATATGCTCAACAGAACTCATCTCTATGAGCGCTGGGAACATCA 1140
QY 381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
DB 1141 GACCATGTTGTTGTTATTAATGATATGCTTACGTAAGGGAACAGAGAGCATGGAT 1200
QY 401 GluTyrThrSerGluIlePheMetGlyGlyLysAsnThrIleValIleuHisAsnThrCys 420
DB 1201 GAGTACACTTCAGAGATATTCATGAGGTGGAAGAGCACCATTTGTTGGACAAACATGC 1260
QY 421 GluAspSerLeuLeuAlaIleProIleIleuAspLeuValLeuLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTCTTAAGCTGCTCTTATTAATCTTGAAGCTTGGCTTCTTGGTGAAGCTGAC 1320
QY 441 ThrArgIleGlnPheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAla 460
DB 1321 ACTGAATCGAGTTTAAAGCTGAAGAAATGAGGAAATTCACATCTCCACCCAGTGTCT 1380
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyLysProValIleAsn 480
DB 1381 ACCATCTTCACTTACCTTCAACAGGCTCTCTGTTCCACGGGTCAACCATGCTGAT 1440
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
DB 1441 GCATTGTCAAAAGCAGCGCAATGCTGGAAGAAACATAATGAGGCTTGTGTGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
DB 1501 CCAGAGATTAACATGATTCCTGAGTACAAAG 1530

```

RESULT 8
ADS81999 standard; cDNA; 1533 BP.

ADS81999;
AC ADS81999;
AC XX
DT 18-NOV-2004 (first entry)
DT XX
XX XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 3.
XX XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX XX
OS Glycine max; line 29004J7P01.
XX XX
FH Key
FT CDS 1..1533
FT Location/Qualifiers
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"

US2003074685-A1.
XX XX
XX 17-APR-2003.
PD 11-MAR-2002; 2002US-00025003.
PF 08-APR-1997; 97US-00835751.
PR

PR	07-APR-1998;	98MO-US006822
XX		
PA	(HITZ/) HITZ M D.	
PA	(SEBA/) SEBASTIAN S A.	
XX		
PI	Hitz WD, Sebastian SA;	
XX		
DR	WPI; 2004-639957/62.	
PR	P-PSDB; ADS82000.	

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, PT phytic acid and inorganic phosphate content of soybean seeds.
XX
Example 8; SEQ ID NO 9; 34pp; English.
XS

Example 8; SEQ ID NO 9; 34pp; English.

The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol 1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytide acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homologous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.

SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	2.65e-251
Score:	2593.00
Percent Similarity:	99.22%
Best Local Similarity:	98.43%
Query Match:	98.56%
DB:	13
Gaps:	0
Length:	1533
Matches:	502
Conservative:	4
Mismatches:	4
Indels:	0

US-10-718-952-12 (1-510) X ADS81999 (1-1533)

QY	1	MetPhelIeIuEnpHeIySeValGUseProAnValLveTyrThcIuNrhcIuIle	20
Db	1	AtGTCATCGAGAAATTTAAAGTTGAGTGCTCTTAATGTAAAGTACCCGAGCTAGAGNT	60
QY	21	GInSeValTyRAnTyrgIuNrHgIuLeuValHsGIuAsnARgAnGIyThrTyR	40
Db	61	CAGCCGTGACAACTTCGAAACCAACCGAACTTGCACAGAGACGGAATGGCACTTAT	120
QY	41	GIuTrIpleIaIySPoLySeSeValAsnTyRGInPhcIyThRAsnThRHisValPro	60
Db	121	CAGTGGATTTGTCAAAACCCAAATCTGTGCMAATTAAGAAATTTAAACCAACTCATGTCTCT	180
QY	61	LysIeGIuIyValMetIeUValGIyTrIpGIyGIyAsnAnGIySeThRIduNrHgIyGIy	80

Db	181	AAATTAGGGGTAATGCTTGTGGGTTGGGGGTGGAAACAACGGCTCAACCCCTACCGGTGT	240
QY	81	VallIleAlaAnaArgIuAspIleSerTrpAlaThrIysAspLysIleGlnIleAlaAsn	100
Db	241	GTTATTGTGTAACCGAGAGGGCATTTTCATGGGCTCAAAAGACAGAAATTCACAAAGCCAAAT	300
QY	101	TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu	120
Db	301	TACTTTGGCTCCCTCACCCCAAGCCTCAGCTATCCGAGTTGGGTCTTCCAGAGGAGAGAA	360
QY	121	IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly	140
Db	361	ATCTATGCCCATTCAGAGGCTGCTCTTCATAGGTATACCTTGACGACATTTGTGTGGG	420
QY	141	GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp	160
Db	421	GGATGGGATATCACCAACATGTAACCTGGCTGATGCCATGGCCAGGGCAAAAGTGTGGAC	480
QY	161	IleAspLeuGlnLysGlnLeuArgProTyrMetGlnSerMetValProLeuProGlyTle	180
Db	481	ATCGATTTGGCAGAAAGCAGTTGAGGCTTTACATGGAATCCATGCTTCACCTCCGGATTC	540
QY	181	TyrAspProAspPheIleAlaAlaAsnGlnGlnIuArgAlaAsnAsnValIleLysGly	200
Db	541	TATGACCCGGATTCATATGCTGTGCCAACCAAGAGAGGTGCCCAACAGTATCAAGGCG	600
QY	201	ThrLysGlnGlnIleValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr	220
Db	601	ACAAAGCAAGACAGATTCACCAAAATATCTAAAGACATCAAGGGCTTTAAGAGGCCACC	660
QY	221	LysValAspLysValValIleuTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal	240
Db	661	AAAGTGCAAGAGGTGGTGTCTGTGGACCTGCCAACAGAGAGGTACAGTAATTTGGTT	720
QY	241	ValGlyLeuAsnAspTrpMetGlnAsnLeuLeuAlaValAspArgAsnGluAlu	260
Db	721	GTGGGCTCTTAATGACCCATGGAGAAATCTTGGCTGTGTGGACAGAAATGAGGCTGAG	780
QY	261	IleSerProSerTrpLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn	280
Db	781	ATTTCCTCTTCCACTTGATGATGCATTCCTGTGTGTATGAAAAATGTTCTTTCATTAT	840
QY	281	GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr	300
Db	841	GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATGTGATCTTGGCATCGCAGAGAACT	900
QY	301	LeuIleGlyIleAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp	320
Db	901	TTGATTTGTGTGAGATGATCTTCMAAGGTGTGACCAAAAGAAATCTGTGTGTGAT	960
QY	321	PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn	340
Db	961	TTCTCTGGGGGGGCTGGTATCAAGCCACATCTATAGTCACTTACAAACATCTGGAAAC	1020
QY	341	AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer	360
Db	1021	AATGATGTTAAGAAATCTTTGGCTCCACAAACTTCCGTTCCAAAGAAATCTCCAAAGGC	1080
QY	361	AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGlnHisPro	380
Db	1081	AACGTGTGTGATAGTATAGTCCMAAGCAAAATCCATCTCATAGGCGCTGTGAACATCCA	1140
QY	381	AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp	400
Db	1141	GACCATGTTGTATTATTAAGTATAGTCCCTTACSTTGGGAGACAGCAAGAGCCATGAT	1200
QY	401	GlnTyrThrSerGlnIlePheMetGlyIleLysAsnThrIleValLeuHisAsnThrCys	420
Db	1201	GAGTACACTTCAGAGATATTCATGCGGTGGTGAAGAAGACCAATTTGTTTGGACAAACATGC	1260
QY	421	GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGlnLeuSer	440
Db	1261	GAGGATTCCTCTTAAGTGTCTCTTAATATCTTGGACATTTGCTCTTCTTGCTGACCTCAGC	1320

QY	21	GlnSerVal1TyrAsnThrCgluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr	40
Db	114	CAGTCGCGTGTAACTACGAAACCAACCGGACTTGTTCAGAGAAACGAGATGGCACTTAT	173
QY	41	GlnTPriLeValLysProLysSerValAsnTyrGlnPheLysThrAsnThrHisValPro	60
Db	174	CAGTGGATGTGTCAAAACCAATCTGTGCAATACGAAATTTAAACCAACATCCATGTTCTT	233
QY	61	LysLeuGlyValMetLeuValGlyTyrGlyLysAsnGlySerThrLeuthrGlyGly	80
Db	234	AAATTAGGGGTATGCTTGTGGGTTGGGGTGGAAACAAACGGCTCAACCTCCACGGTGGT	293
QY	81	ValIleAlaAsnArgGluAsnTrpLysSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn	100
Db	294	GTTATATGCTAACCGAGGGGCAATTCATGCGCTTACAAAGAACAGATTCAACAGCAAT	353
QY	101	TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGluGlu	120
Db	354	TACCTTTGGCTCCCTCAACCCAGAGCTCAGCTATCCAGATTGGGTCTTCCAGGAGAGAA	413
QY	121	IleTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly	140
Db	414	ATCTATGCCCCCATTCAGAGCGCTGCTTCCAAAGTGTAAACCTGACGACATGTGTTGGG	473
QY	141	GlyTPAspAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp	160
Db	474	GGATGGGATATACGAAACATGAAACCTGCTGATGCCATGGCCAGGGCCAAAGGCTTTGAC	533
QY	161	IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle	180
Db	534	ATCGATTTGCGAGACAGCTGAGGCTTACATGGAATCCATGCTTCACTCCCCGGAAATC	593
QY	181	TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly	200
Db	594	TATGACCCGGATTTTCATTGCTCTCCAAACCAAGAGGCGTCCAAACAGCTCATCAAGGCG	653
QY	201	ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr	220
Db	654	ACAAAGCAAGACCAAGTTCACAAATCATCAAAAGCATCAAGCGCTTAAAGAAAGCCACC	713
QY	221	LysValAspLysValValLysTrpThrAlaAsnThrGluArgTyrSerAsnLeuVal	240
Db	714	AAAGTGGACAGGTGGTGTACTGTGGACTGCCAACACAGAGGTACGTAATTTGGTT	773
QY	241	ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu	260
Db	774	GTCGGCCTTAATAGACCACTGAGAAATCTCTTGGCGCTGTGTGACAGAAATGAGGTGAG	833
QY	261	IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn	280
Db	834	ATTTCCTCTTCCACCTTGATGATCCATTCGCTGTGTATGAAAAAGTTCCTTCACTTAAT	893
QY	281	GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr	300
Db	894	GGAAACCCCTCAGAACACTTTGTACCAAGGCGATGTGATCTTGCCATCGCGAGAACACT	953
QY	301	LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp	320
Db	954	TTGATGTGGTGAAGATGACTTCAAGAGTGTCAAGACCAAAATGAATCTGTGTGGTGAT	1011
QY	321	PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsn	340
Db	1014	TTCCCTTGGGGGGCTGTATCAAGCAACATCTATATAGTACATTACACCATCTGGGAAC	1073
QY	341	AsnAspGlyMetAsnLeuSerLysProGlnThrPheArgSerLysGluIleSerLysSer	360
Db	1074	AATGATGGTATGAATCTTTCGGCTCCACAAACTTTCCTCCAGAGAAATCTCCAGAAC	1133
QY	361	AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro	380
Db	1134	AAACGTTGTGATGATATGTGTCAACAGCAATGCCATCTCTATAGACCTGTGTAAACATCCA	1193
QY	381	AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp	400

Db	1194	GACCATGTTGTTGTTATTAAGTATGTCCTTACCTAGGAGCAGCAAGAGCCATGGAT	1253
Qy	401	GIuTYrThrSerGIuIlePheMetGIyGIyLysAsnThrIleValLeuHIIsAsnThrCys	420
Db	1254	GAGTACACTTCAGAGATTTTCATCTGGTGGAAGAGCACCATTGTTTGGCAACACATGC	1313
Qy	421	GIuAspSerLeuLeuAlaIaProIleIleLeuAspLeuValLeuLeuAlaGIuLeuSer	440
Db	1314	GAGGATTCCTCCCTTAGCTGCTCCTATATATCTGGACATTCGTGCTTCCTGAGCTCAGC	1373
Qy	441	ThrArgIleGIuPheLysAlaGIuAsnGIuGIyLysPheHisSerPheHisProValAla	460
Db	1374	ACTGAAATCGAGTTTAAAGCTGAAGAAATGAGGGAAATTTCCATCTATTCCACCAGTTGCT	1433
Qy	461	ThrIleLeuSerTYrLeuThrLysAlaProLeuValProProGIyThrProValIaIsn	480
Db	1434	ACCATCTCTCAGCTACCTCACCAGAGCTCCTCTGTTCCACCGGGTACACCACTGTGAAT	1493
Qy	481	AlaLeuSerLysGIuArgAlaMetLeuGIuAsnIleMetArgAlaCysValGIyLeuAla	500
Db	1494	GCATTGTCAAAGCAGCGTGCATCTCGAAAACATATAGAGGGCTTGTTGGATTGGCC	1553
Qy	501	ProGIuAsnMetIleLeuGIuTYrLys	510
Db	1554	CCAGAGAAATACATGATTTCTGAGTACAG	1583
RESULT 11			
ID	AAV62440	standard; cDNA, 1782 BP.	
AC	AAV62440;		
XX			
XX	17-OCT-2003 (revised)		
DT	02-FEB-1999 (first entry)		
XX			
DE	Soybean wild-t-type myo-inositol 1-phosphate synthase cDNA.		
XX			
KM	Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;		
KW	phytic acid; ds.		
XX			
OS	Glycine max; line LR13.		
XX			
FM	Key	Location/Qualifiers	
FT	CDS	54..1586	
FT		/*tag= a	
XX			
FN	W09845448-A1.		
XX			
PD	15-OCT-1998.		
XX			
PF	07-APR-1998; 98MO-US006822.		
XX			
PR	08-APR-1997; 97US-00835751.		
XX			
PA	(DUPO) DU PONT DE NEMOURS & CO E I.		
XX			
PI	Hitz WD, Sebastian SA;		
XX			
DX	WPI; 1998-568353/48.		
DR	P-PSDB; AAW79740.		
XX			
PT	Soybean plants containing altered myo-inositol-1-phosphate gene - useful		
PT	for generating plants with altered levels of e.g. raffinose, stachyose,		
PT	phytic acid, etc.		
XX			
PS	Example 5; Page 44-45; 63pp; English.		
XX			
CC	This is the nucleotide sequence of cDNA encoding the wild-type soybean		
CC	myo-inositol 1-phosphate synthase (MI 1-PS) present in clone p5hm1-1ps		
CC	(ATCC 97970). The clone was isolated from a cDNA library of soybean line		
CC	LR13 by hybridisation to a probe made from MI 1-PS cDNA of Arabidopsis		
CC	thaliana. MI 1-PS is involved in glucose metabolism to phytic acid,		

PD 15-OCT-1998.
XX
XX 07-APR-1998; 98MO-US006822.
XX
XX 08-APR-1997; 97US-00835751.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E. I.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI. 1998-568353/48.
XX
XX P-PSDB; AAW79741.
XX
XX Soybean plants containing altered myo-inositol-1-phosphate gene - useful
XX for generating plants with altered levels of e.g. raffinose, stachyose,
XX phytic acid, etc.
XX
XX Example 5; Page 48-49; 63pp; English.
XX
XX This is the nucleotide sequence of cDNA encoding a mutant soybean myo-
XX inositol-1-phosphate synthase (MI 1-PS) present in clone LR33-10. This
XX clone was isolated from a cDNA library of soybean line LR33 by PCR
XX amplification (see AAV62441-42). Line LR33 was obtained by chemical
XX mutagenesis of wild-type soybean genome and as a reduced raffinose
XX saccharide phenotype. MI 1-PS is involved in glucose metabolism to phytic
XX acid, raffinose and stachyose. Sequencing revealed a single base change
XX mutation (G to T at base 1241) in the LR33 sequence when compared to the
XX wild-type sequence (see AAV62440). The mutation results in a seed
XX phenotype of very low raffinose saccharide sugars, very high sucrose and
XX low phytic acid. The mutated MI 1-PS nucleic acid is used to alter the
XX raffinose saccharide, sucrose, phytic acid and inorganic phosphate
XX content of soybean seeds, leading to useful soybean products, e.g. a seed
XX phytic acid content of less than 14.5 ug/g, a seed content of raffinose and
XX stachyose combined of less than 14.5 ug/g, and a seed sucrose content
XX greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)
XX
XX SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 8 476-251 Length: 1533
XX Score: 2588.00 Matches: 501
XX Percent Similarity: 99.02% Conservative: 4
XX Best Local Similarity: 98.24% Mismatches: 5
XX Query Match: 98.37% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-10-718-952-12 (1-510) x AAV62443 (1-1533)
XX
XX QY 1 MetPheilegluaSnPhelysValgluSerProaenValysTyrrThrgluThrgluile 20
XX Db 1 ATGTTCAATCGAAGATTTTAAAGTTGAGTGCTCTAATGTGAAGTACACCGAGCTGAGATT 60
XX QY 21 GlnservaTYrAsnTYrGluThrThrgluLeuValHlegluaSnArGAsnGlyThrTYr 40
XX Db 61 CAGTCGGGTGACACGTACGAAACCAACCGAACTGTTCAGAGAAACAGGAATGGCACTAT 120
XX QY 41 GlnTPriLevalysProlysSerValAsnTYrGlnPhelysThrAsnThrHisValPro 60
XX Db 121 CAGTGATTTGTAACCAATCTGTCAAAATCGAAATTTTAAACCAACATCTCAATGTTCTT 180
XX QY 61 LysLeuGlYValMetLeuValglYTrpGlYglYAsnAsnGlySerThrLeuThrglyGly 80
XX Db 181 AAATTAGGGGTATATGTTGTGGGTGGGTGGAACAAAGGCTCAACCTTCACCGGTGGT 240
XX QY 81 ValIleAlaAsnArGluAspIleSerTrpAlaThrLYsAspLYsIleGlnGlnAlaAsn 100
XX Db 241 GTTATTGTCTAAACCGAGAGGCAATTCATGGCTACAAAGACAAGATTCAAACAGCAAT 300
XX QY 101 TyrPheGlySerLeuThrglnAlaSerAlaIleArgValGlySerPheGlnGlyGlnGlu 120
XX Db 301 TACTTTGGCTCCCTCAACCAAGCTTCAGCTATCCAGTGGGTCTCTCCAGGAGAGGAA 360
XX QY 121 IleTYrAlaProPheIysSerLeuLeuProMetValAsnProAspIleValPheGly 140

Db 361 ATCTATGCCCATTCAGAGAGCTGCTTCCATATGTTAACTCTGACGACATGTGTTGGG 420
QY 141 GLYTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLYsValPheAsp 160
Db 421 GGATGGATATCAGCAACATATAACTGGCTGATCCCATGGCCAGGCAAGAGTGTTCAC 480
QY 161 IleAspLeuGlnLYsGlnLeuArgProTYrMetGluSerMetValProLeuProGlyIle 180
Db 481 ATTCATTTGCAGAGACGAGTTGAGGCTTTACATGAAATTCATGATCTTCACCTCCCGAATC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluGlnArgAlaAsnValIleLYsGly 200
Db 541 TATAGCCGATTTTCATTTGCTGCAACCAAGAGAGGCTGCAACAGTATCAAGGCG 600
QY 201 ThrLYsGlnGlnGlnValGlnGlnIleIleLYsAspIleLYsAlaPheLYsGluAlaThr 220
Db 601 ACAAGCAAGAGCAAGTTCAACAAATCATCAAAAGACATCAAGGCGTTTAAAGAAAGCCACC 660
QY 221 LysValAspLYsValValLeuTrpThrAlaAsnThrgluArgTYrSerAsnLeuVal 240
Db 661 AAAGTGCAAGGTGGTGTGACTGTGACCTGCCAACAGAGAGGTACAGTAATTTGGTT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValaAspArgAsnGluAlaGlu 260
Db 721 GTGGGCTTAATGACACCATGAGAAATCTTGGACTGTGGACAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTYrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 781 ATTTCTCTTCACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 841 GGAAGCCCTCAGAACACTTTTGTACAGGCGTGAATTGATCTTGCATGCGAGAAACACT 900
QY 301 LeuIleGlyIleAspAspPheLYsSerGlyGlnThrLYsMetLYsSerValLeuValAsp 320
Db 901 TTGATTGGTGAATGATGATCTTCAGAGTGTGACACCAAAATGAAATCTGTGTGGTGTAT 960
QY 321 PheLeuValGlyValAlaGlyIleLYsProThrSerIleValSerTYrAsnHileGlyAsn 340
Db 961 TTCTTGTGGGGCTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLYsGluIleSerLYsSer 360
Db 1021 AATGATGGTATGAAATCTTCCGCTCCCAAACTTCCCTCCCAAGAAATCTCCCAAGAGC 1080
QY 361 AsnValValaAspAspMetValAsnSerAsnAlaIleLeuTYrGlnProGlyGluHilePro 380
Db 1081 AACGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 381 AspHisValValaIleLYsTYrValProTYrValGlyAspSerLYsArgAlaMetAsp 400
Db 1141 GACCATGTGTGTATTAAGTATGCTCTTACGTAGGGGACAGCAATAGGCCATGAT 1200
QY 401 GluTYrTrpSerGluIlePheMetGlyLYsAsnThrIleValLeuHileAsnThrCys 420
Db 1201 GAGTATCATTCAGAGATATTCATGGGTGGAAGAGCAATGTTTSCAACAACAGATC 1260
QY 421 GluAspSerLeuLeuAlaIleProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1261 GAGGATTCCTCTTAAAGCTCTTATTAATCTTGGACTTGTGCTCTTGTGCTGAGCTGAGC 1320
QY 441 ThrArgIleGlnPhelysAlaGluAsnGluGlyLYsPheHisSerPheHisProValAla 460
Db 1321 ACTAGAAATCGAGTTTAAAGCTGAAATAGGAAATTCATCTCATCTCCACCCAGTTCT 1380
QY 461 ThrIleLeuSerTYrLeuThrLYsAlaProLeuValProProGlyThrProValValAsn 480
Db 1381 ACCATCTCAGACTACCTACCAAGGCTCTGTGTCCACGGGTGACCAAGTGTGAT 1440
QY 481 AlaLeuSerLYsGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500

DB 1441 GCATTGTCAAGACGGCTGCAATGCTGAAACATATGAGGCTTGTTGGATTGGCC 1500
QY 501 ProGUAsnAsnMetIleuGluTyrTyr 510
DB 1501 CCAGAGAAATACATGATGTTCTCGAGTACAAAG 1530
RESULT 13
ADQ14494
ID ADQ14494 standard; cDNA; 1533 BP.
XX ADQ14494;
XX ADQ14494;
XX ADQ14494 (first entry)
XX 23-SEP-2004
DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #1.
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
XX myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate; mutant.
XX Glycine max.
XX Synthetic.
XX OS
XX Key Location/Qualifiers
XX CDS 1..1533
XX FT /*tag= a
XX FT /product= "Mutant soybean myo-inositol 1-phosphate
XX FT synthase #1"
XX FT replace(1241,G)
XX mutation /*tag= b
XX US2004128713-A1.
XX 01-JUL-2004.
XX 21-NOV-2003; 2003US-00718952.
XX 08-APR-1997; 97US-00835751.
XX PR 07-APR-1998; 98WO-US006882.
XX PR 26-APR-1999; 99US-00293315.
XX PR 11-MAR-2002; 2002US-00025003.
XX PA (HIRTZ/) HIRTZ W D.
XX PA (SEBA/) SEBASTIAN S A.
XX PA (GRAC/) GRACE D J.
XX PA (STRE/) STREIT L G.
XX Hirtz WD, Sebastian SA, Grace DJ, Streit LG;
XX MPI: 2004-53335/51.
XX P-PSDB; ADQ14495.
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX Claim 10; SEQ ID NO 5; 48bp; English.
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with

CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 8,47e-251 Length: 1533
Score: 2588.00 Matches: 501
Percent Similarity: 99.02% Conservative: 4
Best Local Similarity: 98.24% Mismatches: 5
Query Match: 98.37% Indels: 0
DB: 12 Gaps: 0
US-10-718-952-12 (1-510) x ADQ14494 (1-1533)
QY 1 MetPheIleGluAsnPheLeuValGluSerProAsnValTyrThrGluThrGluIle 20
DB 1 ATGTTTCATCGAATTTTATAGCTTGAAGTCTCTAATGTGAATACCGAGACTGAGATT 60
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHleGluAsnArgAsnGlyTyrTyr 40
DB 61 CAGTCGGGTACCACTACGAAACCAACCGAACTTGTTCAAGAGAACAGAAATGGACCTAT 120
QY 41 GlnTrrPilleValLeuProLeuSerValAsnTyrGlnPheLeuThrAsnThrIleValPro 60
DB 121 CAGTGGATTGTCMAACCCCAATCTGTCAAAATGCAATTTAAACCAACATCTCATTTTCT 180
QY 61 LysLeuGlyValMetLeuValGlyTrrPglYAsnAsnGlySerThrLeuThrGlyGly 80
DB 181 AAATTAGGGGTATATCTTGTGGGTGGGTGGAAACCAACGCTCAACCTTCAACGGTGT 240
QY 81 ValIleAlaAsnArgGluAspIleSerTrrPalThrLysAspLysIleGlnIleAlaAsn 100
DB 241 GTTATTGCTTAACCGAGAGGCACTTTCATGGGCTTACAAAGACAAAGATTCACAAAGCAAT 300
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleAlaGlyValGlySerPheGlnGlyGlu 120
DB 301 TACTTGTGCTCCCTCACCCCAACCCCTCACCTTCCAGATGGTCCCTCCAGAGAGAGAA 360
QY 121 IleTrrAlaProPheLeuSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 361 ATCTATGCCCACTTCAAGAGCTGCTCTCCATGTTAACTTCAAGCACTTGTGTTGGG 420
QY 141 GlyTrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 421 GGATGGATATCAGCAACATGAACCTGCTGATGCCATGGCCAGGCAAGGTGTTTGC 480
QY 161 IleAspLeuGlnLysGlnLeuArgProTrrMetGluSerMetValProLeuProGlyIle 180
DB 481 ATCGATTTGCAAGACAGCTGAGGCTTACATGGAATCATGCTTCCCTCCCGGAATTC 540
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
DB 541 TATGACCCGGAATTTATGCTGCAACCAAGGAGCGGTGCAACCAAGCTCATCAAGGGC 600
QY 201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluIleThr 220
DB 601 ACAAAAGCAAGAGCAAGCTTCAACAATATCAAAAGCATCAAGAGGCTTTAAGAAACCAACC 660
QY 221 LysValAspLysValValIleLeuTrrPThrAlaAsnThrGluArgTrrSerAsnLeuVal 240
DB 661 AAAGTGAACAAGGTGGTGTGCTGAGCTGCCAACACAGAGAGGTACAGTAATTTGGT 720
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluIleGlu 260
DB 721 GTGGGCTTAAATGACACCATGAGATCTTGGCTGCTGGAGCAAGAAATGAGGCTGAG 780
QY 261 IleSerProSerThrLeuTrrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
DB 781 ATTCTCTTCCACCTTGTATGCACTTGTGTGTTATGAGAAATGTTCTTTTCAATAT 840

```

QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
DB 841 GGAAGCCCTCAGAACCTTTGTACAGAGGCTGATGATCTTGCCATCCGAGGAACACT 900
QY 301 LeuIleGlyValAspAspPheIleSerGlyIleThrIleMetIleSerValLeuValAsp 320
DB 901 TTGATTGGTGGAGATCACTTCAAGAGTGTCTAGACCCMAAATGAAATCTGTGGTTGAT 960
QY 321 PheLeuValGlyAlaGlyIleIleIleIleIleIleIleIleIleIleIleIleIleIle 340
DB 961 TTCCTTGGGGGCTGTATCAAGCCACATCATATAGTCAGTTACCAACCATCTGGGAAAC 1020
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheAspSerIleGlyIleSerIleSer 360
DB 1021 AATGATGGATGATGATCTTCCGCTCCACAAACTTCCGTTCCAGGAAATCTCCAGAGAC 1080
QY 361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuIleGlyIleProGlyIleIlePro 380
DB 1081 AACGTTGTTGATGATGATGTCTACAGCAATGCCATCTTATAGAGCTGTGATCAATCCA 1140
QY 381 AspHisValValValIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 400
DB 1141 GACCATGTTGTTGTTATTAAGTATGCTTACGTAGGGGACAGCAATAGAGCCATGAT 1200
QY 401 GlyIleThrSerGlyIlePheMetGlyIleIleIleIleIleIleIleIleIleIleIle 420
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAGACACATGTTTGCAACAACATGC 1260
QY 421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleLeuAlaGluLeuSer 440
DB 1261 GAGGATTCCTTCTTACCTCTCTATTTCTTGGACTTGGCTTCTTCTGCTGACTCAC 1320
QY 441 ThrArgIleGlnPheIleValAspGluIleIleIleIleIleIleIleIleIleIleIle 460
DB 1321 ACTAGATGAGATTAAGCTGAAATAGGAGAAATTCACATCACTCCAGCCAGTTGCT 1380
QY 461 ThrIleLeuSerIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 480
DB 1381 ACCATCTCCAGCTACCTCCACCAAGGCTCTCTGTTCCACCGGTACACCACTGGTGAT 1440
QY 481 AlaLeuSerIleGlnArgIleMetLeuGluIleIleIleIleIleIleIleIleIleIle 500
DB 1441 GCAATGTCAGAGCGGTGCAATGCTGAAACATATATAGAGGCTTGTGATTGGCC 1500
QY 501 ProGluAsnAsnMetIleLeuGlyIleIleIleIleIleIleIleIleIleIleIle 510
DB 1501 CCAGAGAAATACATGATTTCTCGAGTACAAAG 1530
RESULT 14
ADS81997
ID ADS81997 standard; cDNA, 1533 BP.
XX
AC ADS81997;
XX
DT 18-NOV-2004 (first entry)
DE Soybean myo-inositol 1-phosphate synthase cDNA mutant #1.
XX
KM Soybean plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX rafinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
OS Glycine max; line LR33.
XX
XX Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
FT replace(1188,G)
FT /tag= b
XX
XX mutation
XX
XX US2003074685-A1.

```

```

PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
PR
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PI
DR Hitz WD, Sebastian SA;
DR WPI, 2004-639957/62.
DR P-PSDB; ADS81998.
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX
PS Claim 8, SEQ ID NO 5, 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
XX soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX phosphate synthase having decreasing capacity for the synthesis of myo-
XX inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX complement, subfragment or the complement of the subfragment, operably
XX linked to suitable regulatory sequences, where expression of the chimeric
XX gene results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX comprising the chimeric gene (with a heritable phenotype of a seed phytic
XX acid content of less than 14.5 micromol/g, a seed content of raffinose plus
XX stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX greater than 200 micromol/g, provided that the plant is not LR33), seeds
XX from the plant, making a soybean plant with the heritable phenotype
XX (comprising crossing LR33 or the plant comprising the chimeric gene with
XX an elite soybean plant and selecting a progeny plant of the cross of
XX crossing step that has a heritable phenotype as mentioned above), seeds
XX of soybean plant made by the above method, a soy protein product derived
XX from seeds of a soybean plant (homozygous for one or more gene encoding a
XX mutant myo-inositol 1-phosphate synthase having decreased capacity for
XX the synthesis of myo-inositol 1-phosphate, where the gene confers a
XX heritable phenotype as mentioned above), and making or producing a
XX soybean protein product derived from seeds of a soybean plant with a
XX heritable phenotype as mentioned above. The nucleic acid is useful for
XX altering raffinose saccharide, sucrose, phytic acid and inorganic
XX phosphate content of soybean seeds thus leading to valuable and useful
XX oligosaccharides in soy plants (and other legumes) can lead to flatulence
XX when consumed by humans. The present sequence encodes a mutant myo-
XX inositol 1-phosphate synthase.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 8,47e-251 Length: 1533
Score: 2588.00 Matches: 501
Percent Similarity: 99.02% Conservative: 4
Best Local Similarity: 98.24% Mismatches: 5
Query Match: 98.37% Indels: 0
DB: 13 Gaps: 0
US-10-718-952-12 (1-510) x ADS81997 (1-1533)
QY 1 MetPheIleGlnAsnPheIleValGluSerProAsnValIleIleIleIleIleIleIleIleIle 20
DB 1 ATGTTTCATCGAAGATTTTAAGTGTGAGTCTATATGGAAGTACACGGAACCTGAGATT 60
QY 21 GlnSerValIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 40
DB 61 CAGTCCGTTACACTACGAAACCCGAACTTGTTCACGAAACAGGAAATGGCACTTAT 120
QY 41 GlnTrpIleValIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 60

```

```

Db      121 CAGTGGATTGTCAACCCAAATCTGTCAATAGCAATTTAAACCAACATCTTCCT 180
Qy      61 LysLeuGlyValMetLeuValGlyTyrGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db      181 AATTAGGGGGTAAATGCTGTGGGGTGGGAAACAACGGCTCAACCCCTCAACGGTGGT 240
Qy      81 VALLIleAlaAsnArgGluAspIleSerThrPalaThrLysAspLysIleGlnGlnAlaAsn 100
Db      241 GTTATTGCTTAACCGAAGGCAATTTTCATGGCTCAAAAGACAAGATTCAACAGCCAAAT 300
Qy      101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db      301 TACTTTGGCTCCCTCAACCAAGCTCAAGCTATCCAGTTGGGTCTTCCAGGGAGAGAA 360
Qy      121 ILeTyrAlaProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db      361 ATCTATGCCCAATTCAGAGACCTGCTTCCAATGGTTAACCTGACGACATTTGTTGGG 420
Qy      141 GYTTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db      421 GGATGGGATATCAGCAACATGAACTGCTGATGCCATGGCCAGAGGCAAAAGGTGTGAC 480
Qy      161 ILeAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
Db      481 ATCGATTTCAGAGACAGTGGAGGCTTTCATGGAATTCATGCTTCACCTCCCGGAATC 540
Qy      181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLysGly 200
Db      541 TATGACCCGGATTTCATGCTCTGCCAACAAAGAGCGGCCAACAACTCATCAAGGCG 600
Qy      201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysValPheLysGluAlaThr 220
Db      601 ACMAAGCAGAGACCAAGTTCACAAATCATCAAAAGCATCAAGCGCTTTAAGAGAGCAC 660
Qy      221 LysValAspLysValValValLeuTrrPrrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db      661 AATGTCGACAGACAGTGTGTACTGTGAGCTGCCACACAGAGAGGTACGTAATTTGGTT 720
Qy      241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsnGluAlaGlu 260
Db      721 GTGGGCTTAAATGACCACTGAGGAATCTCTTGGCTGCTGTGAGACGAATAGGCTGAG 780
Qy      261 ILeSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db      781 AATTCTCTTCACACCTGTATGCCATTTGCTTGTTATGAAATAGTCTCTTCATTATAT 840
Qy      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db      841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCGTGAATGATCTTCCATCCGAGGAACACT 900
Qy      301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db      901 TTGATTGGTGGAGAGACCTTCAAGAGTGTCAAGCAAAATGAATCTGTCTGTGGTGAAT 960
Qy      321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnIleLysGlyAsn 340
Db      961 TTCTCTTGGGGGCGTGTATCAAGCAACATATATAGTACATTACAAACCATCTGGGAAAC 1020
Qy      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
Db      1021 AATGATGGTATAACTTTTGGCTCCAAACCTTTCCGTTCCAAAGAAATCTCCAAAGAC 1080
Qy      361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db      1081 AACGTTGTTGATGATGTGTCAACAGCAATGCCATCTCTAAGACCTGTGAACATCCA 1140
Qy      381 AspHisValValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db      1141 GACCATGTGTGTATTAATATATATGTGCTTACGTAGGGGACAGCAATATGACCATGAT 1200
Qy      401 GUTYrThrSerGluIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCys 420

```

```

Db      1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAGACACCATTTGTTGCACAACATGC 1260
Qy      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db      1261 GAGGATTCCTCTTATGCTGTCTCTTATTAATCTTGAGACTTGCTCTTGTCTAGCTCAGC 1320
Qy      441 ThrArgIleGlnPheLysAlaGluAsnGlnGlyLysPheHisSerPheHisProValAla 460
Db      1321 ACTAGAAATCGAGTTTAAAGCTGAAATAGAGGAAATTCACCTCAATCCACCTGCTT 1380
Qy      461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyTyrProValValAsn 480
Db      1381 ACCATCTCAGCTACTACCAAGGCTCTCTGTTCCACGGGTACACCACTGCTGAAT 1440
Qy      481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db      1441 GCATTGTCAAAGCAGCGTGCATGCTGGAATAATATGAGGGCTTGTGATTGGATGGCC 1500
Qy      501 ProGluAsnAsnMetIleLeuGluTyrLys 510
Db      1501 CCAAGAAATTAACATGATTTCTCGATPACAAG 1530

RESULT 15
AAK90402
ID      AAK90402 standard; cDNA to mRNA; 1950 BP.
XX
AC      AAK90402;
XX
DT      24-SEP-1999 (first entry)
XX
DE      Nicotiana paniculata INPS encoding cDNA.
XX
KW      Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;
XX      water stress; resistance; ds.
XX
OS      Nicotiana paniculata.
XX
PN      JP1187879-A.
XX
PD      13-JUL-1999.
XX
PF      26-DEC-1997; 97JP-00359773.
XX
PR      26-DEC-1997; 97JP-00359773.
XX
PA      (NIBS ) JAPAN TOBACCO INC.
XX
DR      MPI; 1999-451546/38.
XX
DX      P-PSDB; AAY24477.
XX
PT      New INPS gene derived from Nicotiana genus plant - useful for conferring
XX      resistance to water stress to plants.
PS      Claim 2; Page 6-8; bpp; Japanese.
XX
CC      The present sequence encodes Nicotiana paniculata inositol monophosphate
XX      synthase (INPS), designated NpINPS1. INPS can be used to confer water
XX      stress resistance to a plant
SQ      Sequence 1950 BP; 570 A; 377 C; 425 G; 578 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      9,41e-239      Length:      1950
Score:          2470.00      Matches:      471
Percent Similarity: 96.47%      Conservative: 21
Best Local Similarity: 92.35%      Mismatches: 18
Query Match:      93.88%      Indels:      0
DB:              2      Gaps:      0

US-10-718-952-12 (1-510) x AAK90402 (1-1950)
MetPheIleGlnAsnPheLysValGluSerProAsnValLysTyrThrGluThrGluIle 20

```

Db ATGTTATTGAGAAATTTTAAAGGTGAGAGCCCAAGTTAAAGTACACCCGAAAGTGAATTT 151
Qy GlnSerVal1TyrAsnTyrGluThrGlnLeuVal1HsGluAsnArgAsnGlyThrTyr 40
Db CACTGTGCTATGATTAATCAACCACTAGATTAAGTTCATGATGAGAAATAAGGACATAT 211
Qy GlnTrp1LeuValProLysSerValAsnTyrGlnPheLysThrAsnThrHsValPro 60
Db CAATGAGACCGTCACACCTTAAGCTGTCAAAATATGAGTTCAAGACTGATGATTCATGTTCC 271
Qy LysLeuGlyValMetLeuValGlyTyrGlyValAsnGlnGlySerThrLeuThrGlyGly 80
Db AAATAGGGGTATCTCTGTGATGAGGTGGAACCAATGGTTCACCTTGACCGGTGTGT 331
Qy Val1Leu1AsnArgGluAsp1LeuSer1Phe1AlaThrLysAspLys1LeuGln1Asn 100
Db GTTATGTGTAACAGAGAAAGAAATTTTCATGGGCCACCAAGATAGGTGTCAACACCAAT 391
Qy TyrPheGlySerLeuThrGln1Asn1Ala1LeuVal1GlySerPheGlnGlyGlu 120
Db TACTTTGGCTCTCTTACTACAGGCTTCTACTATTCAGATTTGGGTCTTTCATGAGAAAG 451
Qy 1LeuTyr1AlaProPheLysSerLeuLeuProMetVal1AsnProAspAsp1LeuValPheGly 140
Db ATCTATGCTCCATTTAAAGCCTCCTCCAAATGTCAAATCCAGATGAGGTAGTGTGGA 511
Qy GlyTTPAsp1LeuSerAsnMetAsnLeu1Asp1AlaMet1AlaArg1AlaLeuValPheAsp 160
Db GGATGGGACATGAGCAATGATTAAGCATGATGCAATGCGCAGGAGCTTAAGGTATTTGAT 571
Qy 1LeuAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyLe 180
Db ATTGATCTACAAAGACAGTGAAGCCCTACATGATCATGATGCTCCACTCCCTGGTATC 631
Qy TyrAspProAspPhe1Leu1Ala1AsnGlnGluGluArg1AsnAsnVal1LeuGly 200
Db TATGACCTGATTTTCACTGCTGCTAACCAAGGCTCACCGTCCACACGTCATCAAGGA 691
Qy ThrLysGlnGluGlnValGlnGln1Leu1LeuAsp1LeuValPhe1AspGln1AlaThr 220
Db ACCAAGAAAGAAACAATGATCAATCAATTAAGATATTAGGAGTTTAAGAAAGAAAC 751
Qy LysVal1AspLysVal1Val1LeuTyrThr1AsnThrGluArgTyrSerAsnLeuVal 240
Db AAAGTGAACAAGTGTGATATTGGAAGTCTTAACACTGAAGATACGATATGAGTT 811
Qy Val1GlyLeuAsnAspThrMetGluAsnLeuVal1Ala1Val1AspArgAsnGluArg1 260
Db GTTGGACTTAATGACATATGAAAAACCTCTTGTCTGTGACACAGAAATGAACTGAA 871
Qy 1LeuSerProSerThrLeuTyrAla1Leu1CysVal1MetGluAsnValProPhe1LeuAsn 280
Db ATATCTCTCCACTTGTATGCTATGCTGATCTTGAAATGATGCTTTTATTAAT 931
Qy GlySerProGlnAsnThrPheValProGlyLeu1LeuAspLeuAla1Leu1ArgAsnThr 300
Db GGAAGCCCCCAACAACCTTTGTCAGGCTCATTTGATTTGCCATCAAGAAACACA 991
Qy Leu1LeuGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerVal1LeuValAsp 320
Db TTGATTTGGTGTGATGACCTTAAGAGTGTCAACCAAAATGAAAGTCACTGCTGTGAT 1051
Qy PheLeuVal1Gly1AlaGly1LeuLysProThrSer1Leu1SerTyrAsnHs1LeuGlyAsn 340
Db TTCTGTGTGGAGCTGTGATTAAGCCAAATATGATGCTACACCAATTTGGGTATC 1111
Qy AsnAspGlyMetAsnLeuSer1AlaProGlnThrPheArgSerLysGln1LeuSerLysSer 360
Db AATGATGGAATGATGATCTGTCTGCCCTCAAACTTTCCGTCAAAGAGATCTCGAAAGT 1171
Qy AsnVal1Val1AspAspMetVal1AsnSerAsnAla1LeuTyrGluProGlyGlyHsPro 380
Db AATGTTGTGATGACATGCTTCAAGCAATGCAATCCTTTATGAGCTTGAGAGACCTCT 1231

Qy AspHsVal1Val1LeuLysTyrVal1ProTyrVal1GlyAspSerLysArg1AlaMetAsp 400
Db GACCATGTTGTTGATTAAGTATGTCGCAATGTGGGAGACAGCAAGGGCAATGGAT 1291
Qy GluTyrThrSerGlu1LeuPheMetGlyGlyLysAsnThr1Leu1LeuHs1AsnThrCys 420
Db GAGTACACATCTGAGATTTTCATGGGGGAAACACCAATTTGTTGCACAAATATCTTGT 1351
Qy GluAspSerLeuLeuAla1AlaPro1Leu1LeuAspLeuVal1LeuVal1GluLeuSer 440
Db GAGGATTCACCTTTTACTGCTCCAAATTAATTAATGATTTGGTCTTCTTGTGCAACTCAGT 1411
Qy ThrArg1LeuGlnPheLysAlaGluAsnGluGlyLysPheHs1SerPheHs1ProVal1Ala 460
Db ACCCGCATTCAGCTCAAAAGCTGAAGAGAGGAGGTAAAGTTCACCTCTCCACCCGTGGCT 1471
Qy Thr1LeuSerTyrLeuThrLysAlaProLeuVal1ProProGlyThrProVal1ValAsn 480
Db ACTATCTCAGCTACCTTACCAAGGCTCTCTGTGATCACCAAGGTACACCAAGGTGAT 1531
Qy AlaLeuSerLysGlnArgAlaMetLeuGluAsn1LeuMetArgAlaCysVal1GlyLeuAla 500
Db GCACCTCAAAAGCAAGGAGCATGCTTGAGAAACATATTGAGGCTTGTGTGGACTTCCA 1591
Qy ProGluAsnAsnMet1LeuGluTyrLys 510
Db CCAAGAAACAACATGATTCCTGAATACAA 1621

Search completed: June 7, 2005, 17:48:12
Job time : 671.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 7, 2005, 17:21:22 ; Search time 3731.5 Seconds

(without alignments)
5202.406 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631

Sequence: 1 MFIEFKVESNVKXTETETI.....NIMRACVGLAENNNILYK 510

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O=/cgn2.1/USPTO_spool_p/US10718952/runac_06062005_173400_12856/app_query.fasta_1.1294

-DB=EST -QFMT=fastap -SUFFIX=sp2n.rst -MINMATCH=0.1 -LOOPEXT=0

-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS-human40.cdi -LIST=45

-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10718952 @CGN 1.1.6628 @runac_06062005_173400_12856 -NCPU=6 -ICPU=3

-NO MAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONGLOC

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gse1:*

9: gb_gse2:*

Result No. Score Match Length DB ID Description

1 2356 89.5 1740 9 CL961092 OaIFCC005

2 2077 78.9 2360 1 AJ583520 OaIFCC005

3 1977.5 75.2 1828 3 AX8084MK BX827819 Arabidops

4 1606 61.0 1813 3 AK079323 AK079323 Mus muscu

5 1606 61.0 1814 3 AK005029 AK005029 Mus muscu

6 1561 59.3 988 7 CK279950 CK279950 EST724028

7 1455 55.3 936 7 CK279064 CK279064 EST725142

8 1438 54.7 972 7 CV135801 CV135801 EST847010

9 1437 54.6 1613 3 CR599696 CR599696 full-leng

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

10	1427	54.2	914	7	CK275352	EST721430
11	1422	54.0	953	7	CO414034	EST844419
12	1391	52.9	988	7	CO113870	GR_Eb014
13	1375	52.3	865	7	CO085839	GR_Ea03E
14	1374	52.2	883	6	CD438045	BL01N0508
15	1374	52.2	891	7	CV290142	aofo1-10m
16	1367	52.0	906	7	CK320178	L2P08a02
17	1363	51.8	912	7	CV142333	EST853542
18	1359	51.7	867	6	CB973407	CAB30002
19	1357	51.6	860	7	CO072025	GR_Ea30P
20	1354	51.5	892	7	CO409215	EST839600
21	1347	51.2	813	7	CK768601	Gm-r1030-
22	1341	51.0	842	7	CF211061	CAB20007
23	1333	50.7	891	7	CF212628	CGF10065
24	1331	50.6	904	7	CV148446	EST859655
25	1335	50.4	836	7	CF518404	CAP007_1
26	1319	50.1	860	7	CO082095	GR_Ea461
27	1304	49.6	960	7	CK274458	EST720536
28	1303	49.5	841	7	CO075472	GR_Ba36C
29	1285	48.8	830	7	CO107630	GR_Eb003
30	1280.5	48.7	1466	3	CR600268	full-leng
31	1280.5	48.7	1513	3	CR606341	full-leng
32	1273	48.4	831	7	CO132349	GR_EB45C
33	1262.5	48.0	834	7	CK933094	CGF100434
34	1260	47.9	970	7	CV147272	EST858481
35	1254	47.7	816	6	CB975215	CAB30006
36	1249	47.5	785	6	CB682609	OSJNFF10C
37	1248	47.4	878	7	CV137015	EST848224
38	1230	46.8	795	7	CO113284	GR_EB013
39	1228	46.7	800	7	CO112984	GR_EB004
40	1228	46.7	843	7	CK267514	EST713592
41	1222	46.4	739	4	BG124172	EST469818
42	1220	46.4	799	7	CF401745	RTW4_14
43	1209	46.0	783	2	BE130712	L48-918T3
44	1208	45.9	779	6	CD448699	EK07D2305
45	1207	45.9	851	7	CO129883	GR_Eb31F

ALIGNMENTS

RESULT 1
LOCUS CL961092 1740 bp DNA linear GSS 21-SEP-2004
DEFINITION OaIFCC005463 Oryza sativa Expressed Sequence Library Oryza sativa (indica
cultiyar-group) genomic, genomic survey sequence.

ACCESSION
CL961092

VERSION
GSS.

KEYWORDS
SOURCE

ORGANISM
Oryza sativa (indica cultivar-group)

REFERENCE
AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

DEPARTMENT

INSTITUTE

TELEPHONE

FAX

EMAIL

CLASS

LOCATION

ALIGNMENTS

RESULT 1
LOCUS CL961092 1740 bp DNA linear GSS 21-SEP-2004
DEFINITION OaIFCC005463 Oryza sativa Expressed Sequence Library Oryza sativa (indica
cultiyar-group) genomic, genomic survey sequence.

ACCESSION
CL961092

VERSION
GSS.

KEYWORDS
SOURCE

ORGANISM
Oryza sativa (indica cultivar-group)

REFERENCE
AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

DEPARTMENT

INSTITUTE

TELEPHONE

FAX

EMAIL

CLASS

LOCATION

/db xref="taxon:39946"
/clone lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences "

ORIGIN

Alignment Scores:
Pred. No.: 8,646-276 Length: 1740
Score: 2356.00 Matches: 451
Percent Similarity: 94.11% Conservative: 28
Best Local Similarity: 88.61% Mismatches: 30
Query Match: 89.35% Indels: 0
Gaps: 0

US-10-718-952-12 (1-510) x CL961092 (1-1740)

```

Oy 1 MetPheIIeGluAsnPhelYsValGluSerProAsnValYsTyrThrGluThrGlu 20
Db 1 ATGTTTCATCGAGAGCTTCGCGTGGAGAGCCCGACGTCGCTACGCGCGCGAGATC 60
Oy 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHISGluAsnArgAsnGlyThrTyr 40
Db 61 GAGTGGACTACCAATACGACGACGAGCTGGTGCAAGAGCCACGACGCGCCCTCC 120
Oy 41 GlnTrrPleValYsProLYsSerValAsnTyrGlnPhelYsThrAsnThrHisValPro 60
Db 121 CGCTGGTGGTCCGCCCAAGTCCGTCGCTACCACTTCCGACCAACCAACCGTCCC 180
Oy 61 LysLeuGlyValMetLeuValGlyTrrPglYglYAsnAsnGlySerThrLeuThrGly 80
Db 181 AAGCTCGGGGGTATCTCGTGGGGTGGGGCGGCAACCGCTCAACCTGACGCTCGG 240
Oy 81 ValIleAlaAsnArgGluAspIleSerTrrPalATrrLysAspLYsIleGlnGlnAlaAsn 100
Db 241 GTCATGCCCAACAGGAGGAAATCTCATGGCGCACCAAGAGCAAGGTGACGACCAAC 300
Oy 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 301 TACTATGGCTCATCACCAGCGCTCCACCATCAAGGTGCGGAGCTACCAACGGGAGAG 360
Oy 121 IleTrrAlaProPhelYsSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
Db 361 ATCTAAGCGCCCTTCAAGAGCTCTCGCCCATGAGTGAACCTCGATGACCTTGTTCGG 420
Oy 141 GlyTrrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 421 GCGTGGACATTAACAATGAACCTGGCTGATGCTATGACCAAGGCCAAGGTACTTGAAC 480
Oy 161 IleAspLeuGlnLYsGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
Db 481 ATTGATCTGCAGAAACAGCTCAGACCTTACATGAGTGCATGATGCTCTCCCGGCATC 540
Oy 181 TyrAspProAspPheIleAlaAlaAsnGlnGluGluValAsnAsnValIleLysGly 200
Db 541 TATGACCCCGACTTATCGCCCAACCAAGGATCCCGCGCAAAATGTCATCAAGAGGC 600
Oy 201 ThrLYsGlnGluGlnValGlnGlnIleIleLYsAspIleLYsAlaPheLYsGlnAlaThr 220
Db 601 ACCAAGAAAGAGCAATGAGCAATCATCAAGCAATCAAGGAATTCAGAAAGAGGC 660
Oy 221 LysValAspLYsValValValLeuTrrThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db 661 AAAGGAGCAAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720
Oy 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu 260
Db 721 GTTGGGCTCAATGACAAATGAGAACTCTCGGCTGTGTGGAACAAGAACGCGGAG 780
Oy 261 IleSerProSerThrLeuTrrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 781 ATATACCAATCAACCTGATGCCATGTGCTGTGCATGAGAGGTATACCTGTCAATTAAC 840
Oy 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300

```

```

Db 841 GCGAGCTCTCAGAAACACCTTTGTGCTGGGCTGATCGATCTTGCTATTAGAACATCTGC 900
Oy 301 LeuIleGlyLYsAspAspPheLYsSerGlyGlnThrLYsMetLYsSerValLeuValAsp 320
Db 901 CTGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 960
Oy 321 PheLeuValGlyAlaGlyIleLYsProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db 961 TTCCATGTTGGTGGTGGATTAAGCCCACTCAATTTGTCAGTCAACCACTTGGGGAAAT 1020
Oy 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLYsGluIleSerLYsSer 360
Db 1021 AATGATGCGATGAACCTTTCGCACTTCAACATTCCTCATCAGAGAAATGTCCAAGAGC 1080
Oy 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db 1081 AATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Oy 381 AspHisValValIleLYsTyrValProTyrValGlyAspSerLYsArgAlaMetAsp 400
Db 1141 GATCATGTTGTTGATCAATGATGTCGATGTTGGAGACAGACAGGCAATGGAC 1200
Oy 401 GlnTyrThrSerGluIlePheMetGlyGlyLYsAsnTrrIleValLeuHisAsnThrCys 420
Db 1201 GAGTACACCTCAGAGATCTTTCATGGGGGTAAAGCACCATCGTTCGACAAACCTGT 1260
Oy 421 GlnAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1261 GAGGACTACTTCTTGGCCGCCGATCATCTTGAATCTTGCTGCTCTTGGAGCTCAGC 1320
Oy 441 ThrArgIleGlnPhelYsAlaGluAsnGlyLYsPheHisSerPheHisProValAla 460
Db 1321 ACCAGGATTCAGCTGAAGCCGAGGGGAGAGAGAGTTCATTCTTCCATCCAGTGCT 1380
Oy 461 ThrIleLeuSerTyrLeuThrLYsAlaProLeuValProProGlyTrrProValValAsn 480
Db 1381 ACCATCTCGAGCTTCTTCAACCAAGGCACCTTGTCTCTCTGCGACACCGTGGTGAAC 1440
Oy 481 AlaLeuSerLYsGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1441 GCCCTGGCAAGACAGAGGCGCAATGCTTGAGAACATCATGAGGGCTGCTGGCTGGCC 1500
Oy 501 ProGluAsnAsnMetIleLeuGluTyr 509
Db 1501 CCCAGAAACAATGATCTTGAAGTAC 1527

```

RESULT 2
AJ583520/c 2360 bp mRNA linear EST 08-JAN-2004
LOCUS AJ583520 Physcomitrella patens cDNA clone pP001015066
DEFINITION similar to putative myoinositol-1-phosphate synthase, mRNA
sequence.
ACCESSION AJ583520
VERSION AJ583520.1 GI:40781590
KEYWORDS EST.
SOURCE Physcomitrella patens
ORGANISM Physcomitrella patens
REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
1 (bases 1 to 2360)
Hone,A., Egener,T., Lucht,J.M., Holtorf,H., Reinhard,C., Schween,G. and Reski,R.
An improved and highly standardised transformation procedure allows efficient production of single and multiple targeted gene-knockouts in a moss, Physcomitrella patens
Curr. Genet. 44 (6), 339-347 (2004)
COMMENT Contact: Schween G
Plant Biotechnology
University of Freiburg
Sommerstrasse 5, D-70104 Freiburg, Germany.
FEATURES
source 1..2360
/organism="Physcomitrella patens"

```

/mol_type="mRNA"
/db_xref="taxon:3218"
/clone="pp001015066"
/issue_type="protonema"
/clone_1ib="ppProci1"

```

ALIGNMENT SCORES:

```

Pred. No.: 1.6e-241 Length: 2360
Score: 2077.00 Matches: 400
Percent Similarity: 88.61% Conservative: 59
Best Local Similarity: 77.22% Mismatches: 51
Query Match: 78.94% Indels: 8
DB: 1 Gaps: 3

```

US-10-718-952-12 (1-510) x AJ583520 (1-2360)

```

QY 1 MetPheilegluaasphelysvalgluserProaenVallyrThrGluThrGluile 20
DB 2150 ATGTTGTCGACAGCTTTCGTTGAGAGTCTCATGTGAGATACAGAGAGACATGATC 2091
QY 21 GluSerValTYrAsnTYrGluThrThrGluLeuValHsGluAsnArg-----AsnGly 38
DB 2090 AAGTCAGACTATGAGTACCAACACCGAGTTCGTGACGACGACGACGACGATGGC 2031
QY 39 ThrTYrGlnTPileVallyrProlySerValAsnTYrGlnPheleThrAsnThrHs 58
DB 2030 AAGTATGAGTGGTGGCCAGAGCCAGTCCGTCCATACCAATTTGCCCTCAGGCGCAAG 1971
QY 59 ValProllyLeuGlyValMetLeuValGlyTPGlyGlyAsnAsnGlySerThrLeuThr 78
DB 1970 GTGCCTAAGCTGGAGATGATGCTCGTGGGTGGGTCGACACGAGTCCACTCTCAGC 1911
QY 79 GlyGlyValIleAlaAsnArgGluAspIleSerTPAlaThrLysAsp-LysIleGlnI 98
DB 1910 GCTGATATCTCGCCACCAAGAGATATCTCATGGGTGACCAAGATGCGGTGACAGA 1851
QY 98 nAlaAsnTYrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySer-PheGln- 117
DB 1850 CGCCAAATTAATCTGGATCCCTCCTCAGAGCTTGGCCGCGGTGGCTCTCTTCCAGG 1791
QY 118 --GlyGluGluIleTYrAlaPProPheLysSerLeuLeu---ProMetValAsn-ProAsp 135
DB 1790 GTGGAAGAAGATTAATGATCTTTAAAGACATTTCAACCGATGTCACCCCAAC 1731
QY 136 AspIleValPheGlyGlyTYrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArg 155
DB 1730 GATCTCGTATCGCGGCTGGGATATCTCCACATGAACTTGGCCGATGCCATGATCGT 1671
QY 156 AlAlaValPheAspIleAspLeuGlnLysGlnLeuArgProTYrMetGluSerMetVal 175
DB 1670 GCGCCGCGTGTGAGATCATGATCTTACAGAACATTCGCTCATGACAGATTCGGA 1611
QY 176 ProLeuPProGlyIleTYrAspProAspPheIleAlaAsnGlnGluLysArgAlaAsn 195
DB 1610 CCCATGCTGTGATTTACGACCCCACTTCATGCTGCTACACAGAGAGCTGTGCGAAC 1551
QY 196 AsnValIleLysGlyThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAla 215
DB 1550 AATCATCAAGAGGTTCCAGAGAGAGCAATAGAGCGGATTTATCAGACATACGGAAC 1491
QY 216 PheLysGluAlaThrLysValAspLysValValIleuTPThrAlaAsnThrGluArg 235
DB 1490 TTCAAGGGCGCAACCAAGGTGACAGAGTGTGATCTGTGACCGCTAACACAGAGCGT 1431
QY 236 TyrSerAsnLeuValValGlyLeuAsnAspThrMet-GluAsnLeuLeuAlaIleValAs 255
DB 1430 TACAGCGAGTATGTGGGTGTGATGACACCAAGAGAAATGTTGATTCCTGGA 1371
QY 255 pArgAsnGluLysGluLysSerProSerThrLeuTYrAlaIleAlaCysValMetGluAs 275
DB 1370 GAAGGAGAAATTAATCTGCGCTGACCTTGTATGCTCGCTTGCATCAGAGACA 1311

```

```

QY 275 nValProPheIleAsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAl 295
DB 1310 GATTCCTTCATCAACGGCAGTCCCAAGACATTCGCTCGAGTATTAATGAATGCC 1251
QY 295 alIleAlaArgAsnThrIleuIleGlyLysAspAspPheLysSerGlyGlnThrLysMetL 315
DB 1250 TGTGAAAAGAACTCATTAATGTTGTGTGACGACTTCAAGACGGGTCAAAACAGATGAA 1191
QY 315 sSerValLeuValAspPheLeuValGlyAlaGlyLysProThrSerIleValSerTy 335
DB 1190 ATCCGTTCTAGAGACTTCTTATGTTGAGCGGGAATTAAGCAACGTCATATTGTAAGTA 1131
QY 335 rAsnHsIleuGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerTy 355
DB 1130 CAACCATTTGGGCAACAGATGATGATGACTTTCGCGCCCTCAACATTCCTGCCA 1071
QY 355 sGluIleSerLysSerAsnValValAspAspMetValAsnSerAsnAlaIleLeuTyG 375
DB 1070 GGAGATCTTCCAAAGACATGATGTGACGACATGTTGGCTAGCAATTCATTAACGC 1011
QY 375 uProGlyGluHsPProAspPheValValIleLysTYrValProTYrValGlyAspSe 395
DB 1010 GCCCGGTGAACATCCCGACCAAGTATTGTGATCAAGTACGCTCTTATGTGGGTGACAG 951
QY 395 rLysArgAlaMetAspGluTYrThrSerGluIlePheMetGlyLysAsnThrIleVal 415
DB 950 CAAGAGGGCAATGACGAGTACATTCGAGAGTTCATGGCGGACGTAACATATCGT 891
QY 415 lIleuHsAsnThrCysGluAspSerLeuLeuAlaIleProlleIleLeuAspLeuValLe 435
DB 890 CATGACATATAGTCGGAAGATCTTTACTGGCCGCGGCTCATCTTGATCTTGCT 831
QY 435 uLeuAlaGluLeuSerThrArgIleGlnPheLysAlaGluAsnGluGlyLysPheHsSe 455
DB 830 GCTGCTAGCTCTGACAGATCAATTCACGAAAGACGACGAGAAAGTTCATTC 771
QY 455 rPheHsProValAlaThrIleLeuSerTYrLeuThrLysAlaPProLeuValProProG 475
DB 770 TTTCACCCGGTGGCTGATCTCTGACTCACTCAAGAGTCTCCCTGTGTAACCCCAAG 711
QY 475 YThrProValValAsnAlaLeuSerLysGlnArgAlaMetLeuGluGluAsnIleMetArgAl 495
DB 710 TACCCCAATGTTTAAACGGTGTGCTTACACAGAGGCGCATGCTAGAGATATCATCGCGC 651
QY 495 aCysValGlyLeuAlaProGluAsnAsnMetIleLeuGluTYrLys 510
DB 650 GTGATGTGGGTGTCCCGGACATTAACATGATGCTAGAAATCAAG 605

```

RESULT 3
CNS044MK 1828 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSURPGH24ZE06 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
BX827819.1 GI:42462602
ACCESSION BX827819.1
VERSION HTG; GSUR_cDNA.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1828)
Castelli,V., Aury,J.M., Jallion,O., Winkler,P., Clepet,C.,
Menard,M., Grnaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1828)
AUTHORS Genoscope.
TITLE Direct Submission

JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT The sequences are based on single pass reads. Life technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castel11 V., Aury J.M., Jailion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.

URCV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length

ftp://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

location/Qualifiers

1.1828

organism="Arabidopsis thaliana"

/mol_type="mRNA"

/strain="COL-0"

/db_xref="taxon:3702"

/clone="GSLRPH24Z06"

/tissue_type="Hormone Treated Callus"

/plasmid="pCMVSPORT 6"

complement(1..1828)

gene "At4g39800"

ORIGIN

Alignment Scores:

Pred. No.: 1.51e-229 Length: 1828

Score: 1977.50 Matches: 397

Percent Similarity: 82.55% Conservative: 43

Best Local Similarity: 74.48% Mismatches: 52

Query Match: 75.16% Indels: 41

Gaps: 6

DB:

US-10-718-952-12 (1-510) x CNS0A4MK (1-1828)

QY 1 MetPheilegluaSnpheylsValgluSerProanVallyrThyThgluThgluile 20

Db ::::: 1977.50

58 ATTTTATTGAGAGCTTCAAGTTGAGAGCCCGAACGTGAATGACAGAGATGAGATT 117

QY 21 GluSerValTYrAnTYrGluThrThrgluValH:sgluAsnArg---AsnGlyThr 39

Db ||||| 118 CACTCTGTCACGATTACGAGACGAGGTTGTCATGAGAACACCGTCAATGCGACT 177

QY 40 TYrGlnTPrlleVallyeProlySserValAsnTYrGlnPheylsThrAsnThrh:sva 59

Db 178 TACCAATGAGATTGTGACGACCAAGACGTGCAAAATGCGATTTCAAAACCGATTGCTGTC 237

QY 60 ProlySleuGlyValMetLeuValglYTrpglyglYAsnAsnGlySerThrleuThrgly 79

Db 238 CCCCAATTAGGGGTTATCTGTGGGTTTGGAGGAGAAACAATGATCAACTCTCACCGCT 297

QY 80 GlyValIleAlaAsnArgGluAspIleSertTPAlaThrlyAspIleleGlnGlnAla 99

Db 298 GGTGTCATTGCCAACAAAGAGAAATCTGTGGGACGACGAGCAAAAGTGCACAAAGCG 357

QY 100 AsnTYrPheglYSerleuThrglnAlaSerAlaIleArgValglYSerPheGlnGlylu 119

Db 358 AATTACTTCGGGTCACTAATCACTCAATTCGATTCGTCGATCTTTTAAAGGTGA 417

QY 120 GluIleTYrAlaProPheylsSerleuEupProMetValAsnProAspAspIleValPhe 139

Db 418 GAGATCTATGCTCTTCAAGAGTCTCTTCCAAATGCGAATCCGATGATGTTGTGTT 477

QY 140 GlyGlyTPAspIleSerAsnMetAsnleuAlaAspAlaMetAlaArgAlaValPhe 159

Db 478 GAGGATGGGATATAGTACATGAATTAAGCAATGCGATGCGTGAAGCCAGGGGTCTT 537

QY 160 AspIleAspLeuGlnlyeGlnleuArgProTYrMetGluSerMetValProleuProGly 179

Db 538 GATATCGACTTGACAAACAGCTCAGCTTACATGAGAAATGTGCGCATCTCCCTGG 597

QY 180 IleTYrAspProAspPheileAlaAlaAsnGlnGluThrgAlaAsnValIlely 199

Db 598 ATCTTCATCTGATTTCATATGCTGCTTATCAAGGTCAGCTGCTTAAACAGTATCAA 657

QY 200 GlyThrgSnglnGlnValGlnGlnIleIlelySAspIlelyValaPheylsGluAla 219

Db 658 GGTACCAAGAGAAACAAGTGCACATCATCAAGACATGAGGAGGATTAAGAGAGAG 717

QY 220 ThrySValaAspIlyValaValleuThrPhrAlaSnThrgAlaUAGTYSerAsnleu 239

Db 718 AATAGAGGTAAAGTTGTGCTCTGACGCGCTAACAGAGCGCTTAAGCAATGTG 777

QY 240 ValValGlyLeuAsnAspThrMetGluAsnleuAlaAlaValaAspArgAsnGluAla 259

Db 778 GTTCGTTGGATGAACGATACATGAGAAATTCATGAGTGTGCGATAGGAGTAGGCT 837

QY 260 GluIleSerProSerThrleuTYrAlaIleAlaCySValMetGluAsnValProPheIle 279

Db 838 GAGATCTCTCTTCAACCTTATGCGATTCCTGTGTTCTTGAAGGAAATCCGTTCAATC 897

QY 280 AsnGlySerProGlnAsnThrPheValProGlyleuIleAspIleAlaIleAlaArgAsn 299

Db 898 AATGAGAGCCCTCAAAACCTTTGTTCCAGTCTAATGATTAAGCGATGAGAACAT 957

QY 300 ThrleuIleGlyGlyAspAspPheylsSerGlyInThrlySmetlySerValleuVal 319

Db 958 GTTTTGATCGGTGAGATGATTTTCAAGAGTGTCAACCAAGATGAATCTGTCTGTT 1017

QY 320 AspPheleuValGlyAlaGlyIlelyeProThrSerIleValSerTYrAsnHileuGly 339

Db 1018 GATTTCCTAGTTGGGCGAGAAATCAAGCTCACTCAATTGTGAGCTCAATCACTTGGA 1077

QY 340 AsnAsnAspGlyMetAsnleuSerAlaProGlnThrPheArgSerlySglnIleSerly 359

Db 1078 AACACAGATGGAATGAATCTCTCAGCTCCACAGACTTTCAGATCTTAAGAGATCTCCAA 1137

QY 360 SerAsnValValaAspAspMetValAsnSerAsnAlaIleleuTYrGluProGlyGlnH: 379

Db 1138 AGCATATGCTGACGATATATGTTGCTAGCAATGATATCTCTTGAAGCTCGGGAAAC 1197

QY 380 -ProAsp-HisValValaValIle-----lySYrValProTYrValGlyAspSerly 397

Db 1198 CTGGACCCATGTTGTTGTCTATCAACAGAGAAACAGTGCACATCATCAAGACATGA 1257

QY 397 TG-----AlaMetArg 401

Db 1258 GGGAGTTTAAAGAGAAATAGGTGATTAAGTTGTGCTCTGACGCGCTTAACACAG 1317

QY 401 luyTYrThrSerGluIlePheMetGlyGlyAsnThrIle-----ValleuHisAsnT 419

Db 1318 AGCGTTATAGCAATGTGCTGTGGATGAACGATACATGAGAAATTCATGAGAGTCTG 1377

QY 419 hrCySgluAspSerleuAlaIleProIleIleleuAspLeuValleuAlaGlu 439

Db 1378 TGATAGGAGTACG-----GCTGAGA 1398

QY 439 euserThrArgIleGlnPheylsAlaGluAsnGlyGlyAsnPheH:SerPheH:SP:ov 459

Db 1399 TCTCTCTTCAACACTTAT-----GGAAAGTTTCACTCTTCACTCTCTG 1443

QY 459 AlaIleThrIleleuSerTYrleuThrlySAlaProleuValProProGlyThrProVal 479

Db 1444 TGCCACCATCTACGACTCACTCAACAGGACCCCTTGTCACGCGGGAATCTCGGCTGA 1503

QY 479 AlaAlaLeuSerlySglnArgAlaMetleuGluAsnIleMetArgAlaCySValGlyL 499

Db 1504 TTAAGAACTGTCAAGACAGCGGCTATGCTGAGAAATATTAAGAGGCGTGGTGAAC 1563

QY 499 euAlaProGluAsnAsnMetIleleuGlnTYrlyS 510

Db 1564 TGCTCCGAGAAACAATGATCATGAAATCAAG 1598

RESULT 4	AK079323	1813 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK079323				
DEFINITION	Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630023013 product:MYO-INOSITOL 1-PHOSPHATE SYNTHASE A1 (EC 5.5.1.4) homolog [Homo sapiens], full insert sequence.				
ACCESSION	AK079323				
VERSION	AK079323.1	GI:26347916			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1				
TITLE	Carninci, P. and Hayashizaki, Y				
JOURNAL	High-efficiency full-length cDNA cloning				
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)				
PUBMED	99279253				
REFERENCE	10349636				
AUTHORS	2				
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,				
JOURNAL	Itoh, M., Komu, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
MEDLINE	Normalization and subtraction of cap-trapper-selected cDNAs to				
PUBMED	prepare full-length cDNA libraries for rapid discovery of new genes				
REFERENCE	Genome Res. 10 (10), 1617-1630 (2000)				
AUTHORS	3				
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,				
JOURNAL	Komou, H., Akiyama, J., Nishi, K., Kitsuaki, T., Tashiro, H., Itoh, M.,				
MEDLINE	Sun, L., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,				
PUBMED	Yanamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,				
REFERENCE	Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsumi, M.,				
AUTHORS	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J.,				
TITLE	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.				
JOURNAL	RIKEN integrated sequence analysis (RISA) system--384-format				
MEDLINE	sequencing pipeline with 384 multiplexed sequencer				
PUBMED	Genome Res. 10 (11), 1757-1771 (2000)				
REFERENCE	11042159				
AUTHORS	4				
TITLE	The RIKEN Genome Exploration Research Group Phase II Team and the				
JOURNAL	PANTOM Consortium.				
MEDLINE	Functional annotation of a full-length mouse cDNA collection				
PUBMED	Nature 409, 685-690 (2001)				
REFERENCE	5				
AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research				
TITLE	Group Phase I & II Team.				
JOURNAL	Analysis of the mouse transcriptome based on functional annotation				
MEDLINE	of 60,770 full-length cDNAs				
PUBMED	Nature 420, 563-573 (2002)				
REFERENCE	6 (bases 1 to 1813)				
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,				
TITLE	Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,				
JOURNAL	Hayashizaki, K., Hayatsu, N., Hizumoto, K., Hirooka, T., Hisokane, T.,				
MEDLINE	Hoti, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,				
PUBMED	Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komu, H., Kouda, M.,				
REFERENCE	Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,				
AUTHORS	Nakamura, M., Nishii, K., Nomura, K., Numazaki, R., Ohno, M., Ohashi, N.,				
TITLE	Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N.,				
JOURNAL	Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,				
MEDLINE	Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takai-Akahira, S.,				
PUBMED	Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,				
REFERENCE	Muramatsu, M., and Hayashizaki, Y.				
AUTHORS	Direct Submission				
TITLE	Submitted (16-APR-2002) Yoshinori Hayashizaki, The Institute of				
JOURNAL	Physical and Chemical Research (RIKEN), Laboratory for Genome				
MEDLINE	Exploration Research Group, RIKEN Genomic Sciences Center (GSC),				
PUBMED	RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,				
REFERENCE	Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,				
AUTHORS	URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222,				

COMMENT
 Fax:91-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: <http://genome.gsc.riken.jp/>
 URL: <http://fantom.gsc.riken.jp/>
 Location/Qualifiers

FEATURES
 source
 1. .1813
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:9630023013"
 /db_xref="taxon:10090"
 /clone="9630023013"
 /issue_type="cerebellum"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="16 days neonate"
 50. .1723
 /note="unnamed protein product; MTO-INOSITOL 1-PHOSPHATE
 SYNTHASE A1 (BC 5.5.1.4) homolog (Homo sapiens)
 (SPTRIQ9NP82, evidence: FASTY, 88.4%ID, 100%length,
 match=1668)
 putative"
 /codon_start=1
 /protein_id="BAC37607.1"
 /db_xref="GI:26347917"
 /translation="MEPAAEILIVSDPVVSPETLEAREYRTTRVREGVLRVQVQ
 ARFPRTAROVPRIGVNLVGMGNNGNSTLAAVLARLRITWPTFRGEANYGYS
 TOAGTVNLGLDNGREBVRVPSALPKVAPNDVLPDGDLSILAAAMRAQVLD
 LDEQLMPHMSLRPRPSVYIPEFLAANTADNLIPETRAQQLERKQIRDRSS
 GDLYKIVLTANTATERPCEVPERNDTANLHTIQLGEVSPSLFVAVSILDECA
 LNSNPQVLPVGAELIASORHVFGEVGDPSQTVKSVLVDPLIGSLKTMSTVSYG
 LKSNPGQNLASAPLOFRSKVETKSSVVDMDVSHNVLAVERPDHGVKIVPVGSG
 KALDEYSELMIGGNTLVLTANTCEDSLAAPIMDLVLTLCORVPSCTSDP
 QGHVTVLSLSTFLRAPLVPPGSPVYNALFRQRCIENTIFRACVGLPQNMHLEHH
 EEPGPQIKGEVVAVSPSPCKKEPTPATNGCTGDANGHPQAPTRKSTA"
 1795. .1800
 /note="putative"
 1813
 /note="putative"
 1813

polya_signal
 polyA_site

ORIGIN

Alignment Scores:
 Pred. No.: 3,366-184 Length: 1813
 Score: 1606.00 Matches: 303
 Percent Similarity: 78.17% Conservative: 91
 Best Local Similarity: 60.12% Mismatches: 104
 Query Match: 61.04% Indels: 6
 Gaps: 4

US-10-718-952-12 (1-510) x AKO79323 (1-1813)

D8 8 ValGusePProAsnValLySTYrThgIunThrgIuIeGInSerValYrAsnYrGlu 27
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 74 GTGGATAGCCCGGAGCTGTCTACAGCCCCGAAACATCGAGCGCGCTACGATGCGG 133
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

D9 28 ThTThGluLeuValHISgluAsnHrGsnGlyThrYrGIntPrIleValLyPProLy 47
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 134 ACAACCGCGGTGACGCCGCGAG-----GAGCGCGTGCAGCG-----GTGCACCCGAG 181
 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Dy 48 SerValAsnYrGlnPhelyrThraSnThrHISValPProLySleuGlyValMetLeuVal 67
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 182 GCTACCGCGTTTACCTTCCGACCGCCCGCGCAGGTGCGCCCACTCGGGGTATTTGGTC 241
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Dy 68 GlyTrpGlyGlyAsnAsnGlySerThrIleuThrGlyGlyValIleAlaAsnHrGluAsp 87
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 242 GCGTGGGGGGGGAACAAGCGCTCCAGCGCTGCTGCTGTTCGGCCAATCGGCTGGCG 301
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Dy 88 IlISerTrpAlaThrIlyAsnPlsYlIeGInGlnAlaSerYrPheGlySerLeuThrGln 107
 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 302 CTAACTGGCCGACGCGACAGCTGCGCAAGAGCAACTATATGATGCTTGAACCG 361
 Qy 108 AAsera1a1earyValGly---SerPheginglyglu1leTyra1a1aProphelys 126
 Db 362 GGGGGACCGCTAACTGGGTCTGATGAGAACGGCCGGAGGTGTGTGCTTCACT 421
 Qy 127 Serleu1euprometVala1asnProaspAph1leValPheglyglYTPAsp11eSeran 146
 Db 422 GCGGTGCTAACCATGATGGTGGCCCCCAACGACCTGTGTGATGATGGGATATCTCGTCG 481
 Qy 147 MetAen1eua1aAsp1a1MetA1aArg1a1aVala1aPheAsp11aAsp1e1n1ySg1n 166
 Db 482 CTGAACCTGGCCGACGCTGCGCGCAGGCTCTGAGATCGCGGTCTGAGAGAACG 541
 Qy 167 Leua1gProTyMeG1uSerMetVal1ProleuPProG1y1eTyraSPProaspPhe1le 186
 Db 542 CTGTGGCCCAATGATGAGAGGCTCGCGCTCGCGCCCTCAAGTCTTAATTCATGCTTCACTC 601
 Qy 167 Ala1a1aenG1nG1uArg1a1aasn1eVal11e1ySg1YTh1ySg1nG1nG1nVal 206
 Db 602 GCTGCCAACGACAGACAGCTGCGCAACCTCATCCCTGCGACAGTCCCAACAGATTG 661
 Qy 207 G1nG1n11e11e1yAsp11e1ySa1a1Phe1ySg1u1a1aTh1yVala1aSp1ySa1a1 226
 Db 662 GAGCAAAATCCGAAGGACATTAGAGATTTCGATCCAGTGGCGGATGGATTAAGGTCAATC 721
 Qy 227 Val1eup1Th1a1a1asnTh1y1uArg1YrSerAen1eua1ValG1Yleu1asn1aSPThr 246
 Db 722 CTGCTGTGACCGCCCAATRGAGAGCGCTTTCGCGAGTGTCTCCAGTGTCCGATATACACA 781
 Qy 247 MetG1uasn1eua1a1aVala1aSPArgaenG1u1aG1u1eSerProSerTh1y1e1u 266
 Db 782 GCAGAAACTTGTCTCATCTATCTCAGCTTGCC---CTGAGAGTGTACCCGTCACACTT 838
 Qy 267 Tyra1a1e1a1aCyValMe1u1asnVal1ProPhe11eAsg1ySerProG1n1asnTh1r 286
 Db 839 TTTGCTGTGGCCAGCATCTGAGAGACCTGCGCTTCTCTCAATGATCCCAACAGAACACA 898
 Qy 287 PheVal1ProG1Yleu11eAsp1eua1a1e1a1aArgaenTh1y1e1G1Yg1YAspAsp 306
 Db 899 CTGGTAACCGGTGCTGGAATCTGCGCTTGGACGCCCATGTGTGTGTAGTGTGATATAC 958
 Qy 307 Phe1ySg1yG1nTh1ySMe1yS1e1yS1e1yS1e1yS1e1yS1e1yS1e1yS1e1yS1e1y 326
 Db 959 TTCAGTCAAGGCGACGCTAAGGTCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1018
 Qy 327 11e1ySPProTh1rSer11eValSer1y1r1asn11e1yS1e1yS1e1yS1e1yS1e1yS1e1y 346
 Db 1019 CTCAGAGACCATGCTCATGCTGCTATTAACCACTTGGGCAACAGACGCGGCAACCACTG 1078
 Qy 347 SerA1a1ProG1nTh1rPheArgSer1ySg1u1eSer1yS1e1yS1e1yS1e1yS1e1yS1e1y 366
 Db 1079 TCTGACCGCTGCACTTCCGCTCCAGAGAGTGCACAAAGAGAGTGTGTGACACATG 1138
 Qy 367 Vala1e1Ser1a1a1e1e1y1yG1uPProG1yG1u11eSProAsp11eVala1a11e 386
 Db 1139 GTTCAAGACCAACATGTGCTTACGCGCTGAGAGCGGCGACCACTGTGTGTATC 1198
 Qy 387 1yS1yTyra1ProTyra1G1YAspSer1ySa1a1aMetA1aSPG1uTyThSerG1u1e 406
 Db 1199 AAATATGTGCTTATGTGGCAGACGACGCTGCTTGGAGAGATTAACCTCCCAAGCTG 1258
 Qy 407 PheMetG1Yg1Y1yAsnTh1r11eVal1eua11e1yS1e1yS1e1yS1e1yS1e1yS1e1y 426
 Db 1259 ATGCTGGGTGGACAAACCTTGTGTCTCAATTAACCTGAGAGATTCGCTCTCGGCC 1318
 Qy 427 AlaPro11e1e1e1yAsp1e1uVal1e1eua1aG1u1eSerTh1yArg11eG1n1e1yS1e1y 446
 Db 1319 GGGCCCAATCAAGCTGACCTAGTGTCTGCTCAAGAGCTGTGTCAACGGCTGAGCTTTCGC 1378
 Qy 447 AlaG1uasnG1uG1Y1ySPhe11eSerPhe11eSProVala1a1Th1r11e1e1yS1e1y1e1u 466
 Db 1379 ACAGACTGGAACCCGAGCTTCAAGGCTTCCACACAGAGTGTCTCCCTGCTTGTCTGCTG 1438

Qy 467 Thr1ySa1a1Pro1eua1a1ProProG1YTh1rProVala1a1a1eua1e1yS1e1yG1nArg 486
 Db 1439 TTTAAGCCCGCGCTTGTGCTCCCGCGGACGCTGTAAGTAATGCGCTTCCGCGACGCG 1498
 Qy 487 AlaMet1eua1a1n11eMetArg1a1CySa1a1G1Yleua1a1ProG1u1asn1aSPMet1le 506
 Db 1499 AGCTGTATCGGAATATTTCAGGGGCTTGTGCGGCTTCCCGCCCAACACCAATGCTA 1558
 Qy 507 LeuG1uTy1yS1e1y 510
 Db 1559 TTAGAGCACAG 1570
 RESULT 5
 AK005029
 LOCUS
 DEFINITION
 Mus musculus adult male liver cDNA, RIKEN full-length enriched library, clone:1300017c10 product:MT0-INOSITOL 1-PHOSPHATE SYNTHASE A1 (EC 5.5.1.4) homolog [Homo sapiens], full insert sequence.
 ACCESSION
 AK005029
 VERSION
 AK005029.2 GI:26339702
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS
 1 Carninci, P. and Hayashizaki, Y.
 TITLE
 High-efficiency full-length cDNA cloning
 JOURNAL
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 PUBMED
 99279253
 10349636
 2
 TITLE
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 JOURNAL
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 MEDLINE
 PUBMED
 20499374
 11042159
 REFERENCE
 AUTHORS
 3
 TITLE
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kasaiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
 JOURNAL
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed capillary sequencer
 MEDLINE
 PUBMED
 20530913
 11076861
 4
 TITLE
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 JOURNAL
 Functional annotation of a full-length mouse cDNA collection
 MEDLINE
 PUBMED
 Nature 409, 685-690 (2001)
 5
 TITLE
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 JOURNAL
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 MEDLINE
 PUBMED
 Nature 420, 563-573 (2002)
 REFERENCE
 AUTHORS
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Aizawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, D., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishii, K., Nomura, K., Numazaki, R., Ono, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,

Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

On Dec 10, 2002 this sequence version replaced gi:12836664.
Please visit our web site (http://genome.gsc.riken.jp/) for further
details.

CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN,
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5'-GAGAGAGAGAGCGCCCACTCGAGTTTCTTTTCTTTCVN 3'], cDNA was
prepared by using trihalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence [5'-
GAGAGAGAGAGATCCAGAGCTCAATTATTTATTAACCCCCCCCC 3']. cDNA was
cleaved with XhoI and ScaI. Cloning sites, 5' end: ScaI; 3' end:
XhoI. Host: SOUR.

FEATURES

source

1..1814
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:DB:1300017C10"
/db_xref="taxon:10090"
/clone="1300017C10"
/sex="male"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
51..1724
/note="unamed protein product; MYO-INOSITOL 1-PHOSPHATE
SYNTHASE A1 (EC 5.5.1.4) homolog [Homo sapiens]
(SPTB/OSNPH2, evidence: FASTV, 88.4%ID, 100%length,
match=1668)
putative"

CDS

/codon_start=1
/protein_id="BAB23756.1"
/db_xref="GI:12836665"
/translation="MEPAEILVDSDDVVSPEITBARYEYRTTSRREGVLARQOPR
ATRTPTTRARQVPRIGVNLVWGKNGNGLTTRAVALANRLRTWPRTRKEKANYGSL
TQAGVNLGLDNGREVEVFPFSAALPMVAPNDLVFGWDLSSLNLEARRRQVLDG
LQEOLEPHMESLRPRPSVYIPFIANQPARADNLPGRRAOQLQIARNDIFSSA
GIDKIVLTANTERCEVEVPGRNDTAEMLHTIOLGLEVSPSTLFAVASLDEDAFL
NGSPONTLVPALELASORHVFVGQTKVSVVDLFISGKTMISVAYNH
LGNNGONLSAFLORSKEVTSVVDMDVSHNVLVYAGGERPDHCIVYIKVYPVGS
KRALDPTYSRLMLGNTITVLTNTGDSLLAPIMLDVLTVELCORVSFCTDSDPEP
QGFHTVLSLFLFAPLVPLPGSPVNNALFRORSCTENFRACVGLPQNHMLLSHKM
RRPGIKRGEVATSPLECKKEPTATGTGCGDANGHPQAPTPLKLSA"

ORIGIN

polyA_signal
polyA_site
/note="putative"
1814
/note="putative"
/note="putative"

Alignment Scores:

Pred. No.: 3.36e-184 Length: 1814
Score: 1606.00 Matches: 303
Percent Similarity: 78.17% Conservative: 91
Best Local Similarity: 60.12% Mismatches: 104
Query Match: 61.04% Indels: 6
DB: 3 Gaps: 4

US-10-718-952-12 (1-510) x AK005029 (1-1814)

QY 8 ValGluSerProAsnValIleYrThrGluThrGluIleGlnSerValIleYrAsnYrGlu 27
DB 75 GTGATATGCCCGAGAGCTGCTCTACAGCCCCAAACATTCAGAGCGCGCTACAGATCCGG 134
QY 28 ThrTrGluLeuValJhIleGluAsnArgAsnGlyThrYrGlnTrpIleValYsProIys 47
DB 135 ACAAGCGCGGTGAGCGCGAG-----GGCGCGTGTCTGG-----GTGAGCCGACG 182
QY 48 SerValAsnYrGlnPheIleYsThrAsnThrHisValProIysLeuGlyValMetLeuVal 67
DB 183 GCTAGCGTTCACCTTCGACGCCGCCGCGAGGCGCCGCGCTCGGCTCATGTGTGTC 242
QY 68 GlyTrpGlyGlyAsnAsnGlySerThrLeuThrGlyGlyValIleAlaAsnArgIleuSer 87
DB 243 GGCTGGGGGGAACAACAGCTCCACGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 302
QY 88 IleSerTrpAlaThrIleYsAspIleIleGlnIleAlaAsnYrPheGlySerLeuThrGln 107
DB 303 CTAACTGGCCCAACCGCAGAGCTGCAAGAGGCAACTATATGATCGTTGACCCG 362
QY 108 AlaSerAlaIleArgValGly--SerPheGlnIleGluIleIleYrAlaProPheIys 126
DB 363 GCGGCGACCGGTGAACCTGGGTGTGATGAGAAGCGCGGAGGTGTGTGCTTCAGT 422
QY 127 SerLeuLeuProMetValAsnProAspIleValPheGlyIleYrAspIleSerAsn 146
DB 423 GCGTGTCTACCATGATGTGGCCCCCAACAGCTGGGTGTGTGATGTGGATATCTGTG 482
QY 147 MetAsnLeuAlaAspAlaMetAlaArgAlaYsValPheAspIleAspLeuGlnIleGln 166
DB 483 CTGAACCTGGCCGAGCGCATGTGGCGCGCGCGAGCTCTGAGACTGGCGTCTGAGAGACG 542
QY 167 LeuArgProIleYrMetGluSerMetValProIleuProGlyIleIleYrAspProAspPheIle 186
DB 543 CTGTGGCCCCCATGAGAGAGCTGCGTGGCGCGCGCTGATCATCTCATGATTCATC 602
QY 187 AlaAlaAsnGlnIleGluIleAlaAsnAsnValIleIleGlyThrIleGlnIleGlnVal 206
DB 603 GCTGCCAATCAGACAGCAGCTGGCAACCTCATCTCGGACACGTGCCAACAGTTG 662
QY 207 GlnGlnIleIleIleYsAspIleIleYsAlaPheYsGluAlaThrIleYsValAspYsVal 226
DB 663 GAGCAAAATCCAAAGAGCATTTAGAGATTTCCATTCAGTGGCGGATTTGGATTAAGTCATC 722
QY 227 ValIleuTrpThrAlaAsnThrIleuArgYrSerAsnLeuValIleGlyLeuAsnAspThr 246
DB 723 GTGCTGTGAGCGCCCATATACGAGCGCTTCTGCGAGGTGTCCAGTCCGCAATACACA 782
QY 247 MetGluAsnLeuLeuAlaAlaValAspArgAsnGluIleGluIleSerProSerThrIleu 266
DB 783 GCAGAAATCTGCTCATATCATATCCAGCTTGGC--CTGAGAGTTCACCGTCCACACTT 839
QY 267 TyrAlaIleAlaCysValMetGluAsnValProPheIleAsnGlySerProGlnAsnThr 286
DB 840 TTTCGTTGTGGCCAGCATCTCGAGAGCTGGCGCTTCTCATGATGCCACAGAACACA 899
QY 287 PheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrIleIleGlyYsAspAsp 306
DB 900 CTGGTACCGCGGTGCTGGAACCTGCTTCGAGCGCCCATGTGTTGTATGTGTGATGAC 959
QY 307 PheIysSerGlyGlnThrIleMetIleYsSerValLeuValAspPheLeuValIleGly 326
DB 960 TTCAAGTCAGGGCACAACCTAAGTCAAGTCTGCTGTGTGAGCTTCTCATGCGCTGCGG 1019
QY 327 IleYsProThrSerIleValSerYrAsnHisIleuGlyAsnAsnAspGlyMetAsnLeu 346
DB 1020 CTCAGAGACCATGTCCATGTGATGCTATTAACCTCTGGCAACAGAGCGGAGAACTTG 1079
QY 347 SerAlaProGlnThrPheArgSerIleGluIleSerIysSerAsnValValAspAspMet 366
DB 1080 TCTGCACCGCTGACAGTTCCTCCCAAGAGAGTACAAAGACAGAGTGTGTGACGACATG 1139

QY 367 ValaanserAsnAlaileuTyrgUpProGlyGluHisProAspHisValValle 386
| | | | | : : : : :
Db 1140 GTTCACAGCAACACGCTGCTTACCGCTCGAGAGCGGCCAGACACGCTGTGTGATC 1199
QY 387 LysTYrValProTyValGlyAspSerLysArgAlaMetAspGluTyThrSerGluile 406
| | | | | : : : : :
Db 1200 AAATGTGCTTATGTTGGCCACAGACGCTGCTTGAGAGATACCTCCAGCTG 1259
QY 407 PheMetGlyGlyLysAsnThrIleValleuHisAsnThrCysGluAspSerleuAla 426
| | | | | : : : : :
Db 1260 ATGCTGGGGGAGCAACACCTTGGTGTCCATATATACCTCGAGAGATTGCTCGTGGCC 1319
QY 427 AlaProIleIleLeuAspLeuValleuAlaGluLeuSerThrArgIleGlnPheLys 446
| | | | | : : : : :
Db 1320 GGGCCATCATATCTGACCTAGTGTGCTCACAGAGCTGTGTCAAGCTGTGCTTGC 1379
QY 447 AlaGluAsnGluGlyLysPheHisSerPheHisProValAlaThrIleLeuSerTyrlleu 466
| | | | | : : : : :
Db 1380 ACAGACTCGGACCCGAGGCTTCAGGGCTTCCACACAGTGTCTCCCTGCTTAGCTTCTG 1439
QY 467 ThrIlysaIaProLeuValProProGlyThrProValIValAsnAlaLeuSerLysGlnArg 486
| | | | | : : : : :
Db 1440 TTAAAGCCCGCTGTGCCCCCGGAGAGCCCTGTATGTAATGCTCTTCCGACAGGCC 1499
QY 487 AlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaProGluAsnMetIle 506
| | | | | : : : : :
Db 1500 AGCTGTATCGAGAAATATTTTCAGGCTTGCCTGGGGCTCCGCCACAGAACCAATGTCA 1559
QY 507 LeuGluTYrLys 510
| | | | | : : : : :
Db 1560 TTNGAGCACAG 1571

RESULT 6
LOCUS CK277950 988 bp mRNA linear EST 03-AUG-2004
DEFINITION ESTT24028 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POAB302 5' end, mRNA sequence.
ACCESSION CK277950
VERSION CK277950.1 GI:39834928
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Bukeriyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (bases 1 to 988)
Buell C.R., Hart A., Zismann V., Karameycheva S.A. and Baker B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other ESTs: EST724029
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via <http://genome.arizona.edu/orders/>.
Seq primer: ART TAG GTG ACA CTA TAG.
Location/Qualifiers
1. 988
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAB302"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="PH10B-Tona"
/clone_lib="potato abiotic stress cDNA library"
/note="vector: PCWVSport6.1; Site_1: EcoRI; Site_2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at

following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.:	4,08e-179	Length:	988
Score:	1561.00	Matches:	301
Percent Similarity:	95.44%	Conservative:	13
Best Local Similarity:	91.49%	Mismatches:	15
Query Match:	59.33%	Indels:	1
DB:	7	Gaps:	0

US-10-718-952-12 (1-510) x CK277950 (1-988)

QY 55 ThrAsnThrHisValProLysLeuGlyValMetLeuValGlyTYrGlyGlyAsnAsnGly 74
| | | | | : : : : :
Db 2 ACTGATGTTCATGTTCCAAA-TTAGGGGTATGCTGTGTGATGGGGGAGAACATGTGT 60
QY 75 SerThrLeuThrGlyGlyValIleAlaAsnArgGluAspIleSerThrPalaThrLysAsp 94
| | | | | : : : : :
Db 61 TCAGCTTGAATGAGAGGTGATTATGCTTACAGAGAAGAAATTTCATGGGCTACCAAGAT 120
QY 95 LysIleGlnGlnAlaAsnTYrPheGlySerLeuThrGlnAlaSerAlaIleArgValGly 114
| | | | | : : : : :
Db 121 AAAGTCAGCAAGCAATTAATTAATCTTGCTGTCTTACACAGGCTCTACTAATTCGAGTTGGT 180
QY 115 SerPheGlnGlyGluGluIleTYrAlaProPheLysSerLeuLeuProMetValAspPro 134
| | | | | : : : : :
Db 181 TCCTTCATATGAGAGGAATATATGCCCCCTTAAAGACATATCTCCATGTGCAATCCA 240
QY 135 AspAspIleValPheGlyTYrAspIleSerAsnMetLeuAlaAspAlaMetAla 154
| | | | | : : : : :
Db 241 GATGATGTAGTGTGTTGAGAGATGGGACATCGACAACTGAATTTGGCAGATGCCATGGCA 300
QY 155 ArgAlaLysValPheAspIleAspLeuGlnLysGlnLeuArgProTYrMetGluSerMet 174
| | | | | : : : : :
Db 301 AGGGCTAAAGTTTGTGACATGATCTTCAAAAGCAGCTGAGGCCCTCATGGAATCCATG 360
QY 175 ValProLeuProGlyIleTYrAspProAspPheIleAlaAlaAsnGlnGluArgAla 194
| | | | | : : : : :
Db 361 GTCCACACGCTGGTATCTATGACCCTGATTCATGTGTGCAACCAAGGCTCAGTGGCT 420
QY 195 AsnAsnValIleLysGlyThrLysGlnGluGlnAlaGlnIleIleLysAspIleLys 214
| | | | | : : : : :
Db 421 AACATATGCATTAAGAGAACCAAGAAAGAACAAATTCATCAATCATTAAGATATTAAG 480
QY 215 AlaPheLysGluAlaThrLysValAspLysValValleuTYrPThrAlaAsnThrGlu 234
| | | | | : : : : :
Db 481 GAGTTTAAGAGAACCAACAAAGGTGATGTTGTTGTGAGCTGTAACATCGAA 540
QY 235 ArgTYrSerAsnLeuValValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaVal 254
| | | | | : : : : :
Db 541 AGATACAGCAAGTGGCTGTGGCTTAATATATCCATGTGAAACCTTTGTCTGTG 600
QY 255 AspArgAsnGluAlaGluIleSerProSerThrLeuTYrAlaIleAlaCysValMetGlu 274
| | | | | : : : : :
Db 601 GATGAAATGAGGCTGAATATCTCTTCCACCTTGTATGCTATTTGCTTATTCGAA 660

QY 503 AspaanMet 505
 Db 959 AACCAATG 967

RESULT 9
 CRS99696
 LOCUS
 DEFINITION full-length cDNA clone CS0D1009YP03 of Placenta Cot 25-normalized

ACCESSION CRS99696
 VERSION CRS99696.1 GI:50480503
 KEYWORDS HTC: CNSLT cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1613)
 AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 CONTACT : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 REMARK Faraday Avenue
 2 (bases 1 to 1613)

REFERENCE 2 (bases 1 to 1613)
 AUTHORS Direct Submission
 TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
 JOURNAL - Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.

FEATURES
 source Location/Qualifiers
 1..1613
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1009YP03"
 /tissue_type="Placenta Cot 25-normalized"
 /plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:
 Pred. No.: 1,2e-163 Length: 1613
 Score: 1437.00 Matches: 276
 Percent Similarity: 70.75% Conservative: 82
 Best Local Similarity: 54.55% Mismatches: 88
 Query Match: 54.62% Indels: 60
 Db: 3 Gaps: 4

US-10-718-952-12 (1-510) x CRS99696 (1-1613)

QY 6 PheylValAlGluSerProAsnValLysTyrThrGluThrGluIleGlnSerValTyrAsn 25
 Db 56 TTCCTTCGTGAGAGCCGAGCGTGTCTACGGCCCGAGGCGATGAGCGGCAATACGAG 115

QY 26 TyrGluThrGluLeuValHisGluAsnArgAsnGlyThrTyrGlnTrpIleValLys 45
 Db 116 TACCGAGACGCGCGGTCCAGCCGAG----- 142

QY 46 ProLysSerValAsnTyrGlnPheLysThrAsnThrHisValProLysLeuGlyValMet 65
 Db 142 ----- 142

QY 66 LeuValIGlyTrpGlyGlyAsnAsnGlySerThrLeuThrGlyValIleAlaAsnArg 85
 Db 143 -----GGTGGCGTT----- 151

QY 86 GluAspIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsnTyrPheGlySerLeu 105
 Db ----- 105

Db 152 -----CTCAAGAGGCCCACTACTAGGCTGCTG 181

QY 106 ThrGlnAlaSerAlaIleArgValGly---SerPheGlnGlyGluIleTyrAlaPro 124
 Db 182 ACTCAGCGCGGACCGCTGAGCTGCGCTGACGCGCGAGGCGCAGAGAGTGTTCCTACCC 241

QY 125 PheLysSerLeuLeuProMetValAsnProAspIleValPheGlyGlyTrpAspIle 144
 Db 242 TTCACCGCGGTCTCCCATGTGTGGCCCAACGACCTGTGTGTGATGTGGAGCATCT 301

QY 145 SerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAspIleAspLeuGln 164
 Db 302 TCGTCGCTGAACCTGCGCGAGCGGANTGGCGCGCGGAGAGTCTGTGACTGGGCGCTGACG 361

QY 165 LysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIleTyrAspProAsp 184
 Db 362 GAGCAACTGTGGCCCGACATGATGAGGCGCTGCGCGCGCGCTTCTGTTCATCTCCCA 421

QY 185 PheIleAlaAlaAspGlnGluGluArgAlaAsnValIleLysGlyThrLysGlnGlu 204
 Db 422 TTCATGCGGCGCCACCAAGCGCGCGCGGACCACTCATCTCCAGGCTCGGTGCGCAG 481

QY 205 GlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThrLysValAspLys 224
 Db 482 CAGCTGAGCAGATCCGACGAGGACATCCGAGACTTCGGGTACGCGCGCGGCTGACAAA 541

QY 225 ValValValLeuTrpThrAlaAsnThrGluArgTyrSerAsnLeuValIGlyLeuAsn 244
 Db 542 GTCATAGTGTGTGGACGCGCAACGAGCGCTTCTGTGAGTATTCAGGCTCAAC 601

QY 245 AspThrMetGluAsnLeuLeuAlaAlaAspArgAsnGluAlaGluIleSerProSer 264
 Db 602 GACACAGCGGAGAACCTCTCTGCGCACCATGTGACTCGGT---CTGAGAGTGTGCGCTCC 658

QY 265 ThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsnGlySerProGln 284
 Db 659 ACGCTCTTCCGCGTGGCCAGCATCTCTGAGAGGCTGTGCTTCTCAATGCGTCCGACG 718

QY 285 AsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeuIleGlyGly 304
 Db 719 AACACCTGTGTGCGCGGAGCTCTTGAGCTCGGTGCGCAGCACCGGGTTTGTGTGGCGGA 778

QY 305 AspAspPheLysSerGlyGlnThrLysMetLysSerValIleValAspPheLeuValGly 324
 Db 779 GATGACTTCAAGTACGCGCAGACCAAGTCAAGTCCGCTGTGTGACTTCTCAATTGCGC 838

QY 325 AlaGlyIleLysProThrSerIleValSerTyrAsnHisLeuGlyAsnAspIleYMet 344
 Db 839 TCCGCGCTCAAGACCATCTCATCTGTGATTACAACCACTGGGCAACACATGGGAGG 898

QY 345 AsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSerAsnValAlaAsp 364
 Db 899 AACCTATCGCGCCATTCAGTTCGCTCTAAGAGAGGTCCAGAGCAACGTGGTGGAC 958

QY 365 AspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisProAspHisValAla 384
 Db 959 GACATGTGTGACAGAGCAACCGAGTCTATACGCCCGCGGAAGCCCTGACACATGCGCTG 1018

QY 385 ValIleLysTyrValProTyrValIGlyAspSerLysArgAlaMetAspGlyTyrThrSer 404
 Db 1019 GTCATCAAGTATGTGCTGTACGTGGTGAACGAGCGCGCTGTGATGATTAACCTTCG 1078

QY 405 GluIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrCysGluAspSerLeu 424
 Db 1079 GAGCTGATGCTGGCGGAAACCAACACTGTGTGCAACAACGTGAGAGACTCGCTG 1138

QY 425 LeuAlaAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSerThrArgIleGln 444
 Db 1139 CTGGCGGACCCATCATGTGTGACCTTGTGCTGACCGAGCTGTGCAAGCGGTGACGCGGTGAGC 1198

QY 445 PheLysAlaGluAsnGluGlyLysPheHisSerPheHisProValAlaThrIleLeuSer 464
 Db 1199 TTCGCACTGACATGAGCCCGAGCGGACCTTTCACCCCGTGTCTCTGCTGACG 1258

Qy 465 TyrleuThrlYsAlaProleuValProProGlyThrProValValAsnAlleuSerlys 484
:::|||||
Db 1259 TTCCTTCACAGGGCCCACTAGTGCCTGCCGCCGACCCGGTGGTCAATGCCGTTTCCGC 1318
Qy 485 GlnAtgAlaMetleuGlnAsn11eMeArGa1aCyValGlyLeuAlaProGlnAsnAsn 504
1319 CAGCCACCTGCTGAGAACACTCTCTCAGGGCTGCGTGGGCTCTCCGCCACAGAACAC 1378
Db 505 MetileuGluTyrlys 510
1379 ATGCTCTGGAACACAA 1396
RESULT 10
CK275352 914 bp mRNA linear EST 03-AUG-2004
LOCUS EST721430 potato abiotic stress cDNA library Solanum tuberosum cDNA
DEFINITION clone POADM85 5' end, mRNA sequence.
ACCESSION CK275352
VERSION CK275352.1 GI:39832330
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (bases 1 to 914)
Buell, C.R., Hart, A., Ziemann, V., Karameyheva, S.A. and Baker, B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
JOURNAL Contact: Robin Buell
COMMENT The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Institute can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/.
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers
1..914
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POADM85"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Toni"
/clone_lib="potato abiotic stress cDNA library"
/note="Vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

Pred. No.: 8,26e-163 Length: 914
Score: 1427.00 Matches: 277
Percent Similarity: 97.95% Conservative: 9
Best Local Similarity: 94.86% Mismatches: 6
Query Match: 54.24% Indels: 0
DB: 7 Gaps: 0

US-10-718-952-12 (1-510) x CK275352 (1-914)

Qy 217 LysGlnA1aThrlYsValAspLyValValLeuTrpThra1aAsnThrGluArGlyr 236
|||
Db 1 AAGAGAAACAGCAAGTGCACAAAGTGTGTTTGTGACCTGCCAACACATGAAAGTAC 60
Qy 237 SerAsnLeuValValGlyLeuAsnAspThrMetGlnAsnLeuVala1aValAspArg 256
|||
Db 61 AGCAGTGTGCTGTGGCTTAATGATACCATGAAAAACCTTTGGCTTGTGTGATAGA 120
Qy 257 AsnGlnA1aGlu1ieserProserThrLeuTrpAl1e1aCyValMetGluAsnVal 276
121 AATGAGGCCGAAATATCTCTCTCCACTTGATCTATTGCTGTATTTGAAAATGTG 180
Qy 277 ProPhe11eAsnGlySerProGlnAsnThrPheValProGlyLeu11eAspLeuAl1e 296
181 CCTTTCATCAATGGAAGCCCAACAAACATTTTGTCCAGGCTCATTTGATTTGCCATC 240
Qy 297 AlaArgAsnThrLeu11eGly11aAspAspPhe11eSerGlyGlnThrLysMetLysSer 316
241 AAGAGAAACATTTGATTTGGTGTGATGACTTTAAGAGTGTGCAACCAAGATGAACTCA 300
Db 317 ValLeuValAspPheLeuValGlyValGly11eLysProThrSer11eValSerTyrAsn 336
301 GTGCTGTGTATTTCTCTTGTGGAGCTGTATTAAGCAACATTAATGTAGCTACAA 360
Qy 337 HisLeuGlyAsnAsnAspGlyMetAsnLeuSer11aProGlnThrPheArgSerLysGlu 356
361 CATTTGGTAAATGATGATGAATGAATATCTGCGGCCCTCAGACTTTCCGCTCAAAGAG 420
Qy 357 11eSerLysSerAsnValValAspAspMetValAsnSerAsnAl1eLeuTyrGluPro 376
421 ATCTGAAAACATATGTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 480
Db 377 GlyGlnHisProAspPheValValVal11eLysTyrValProTyrValGlyAspSerLys 396
481 GGAAGAGCACTGTACCATGTTGTGTATTAAGATGACCATATGTTGCTTATGAGCCT 540
Qy 397 ArgAlaMetAspGluTyrTrpSerGlu11ePheMetGlyGlyLysAsnThr11eValLeu 416
541 AGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 417 HisAsnThrCyGluAspSerLeuLeuAla1aPro11e11eLeuAspLeuValLeuLeu 436
601 CACATATCTGTGAGATTTCTTTTGGCTGCTCAATTAATCTTGATTTGGTCTCTTG 660
Qy 437 AlaGluLeuSerTrpArg11eGlnPheLysAlaGluAsnGluGlyLysPheHisSerPhe 456
661 GCTAAGCTACAGACCCCGATTCACCTCAAGCTAAGAGAGGGTAAAGTTGCTCCTCT 720
Db 457 HisProValAlaThrl1eLeuSerTyrLeuThrLysAlaProLeuValProProGlyThr 476
721 CATCCCGTGGGAGCATCTCTAGCTATCTTACCAAGGCTCTCTGATACCAAGATGACA 780
Qy 477 ProValValAsnAlaLeuSerLysGlnArgAlaMetLeuGlnAsn11eMetArgAlaCy 496
781 CCAAGTGTGATGCTCTTTCAAAACAGAGGGCATGCTTGAGAACATATTGAGGGCTTGT 840
Db 497 ValGlyLeuAlaProGluAsnMet11eLeuGlu 508
841 GTTGACTGGCACCAAGAACATCATGATATTGGAA 876
RESULT 11
LOCUS CO414034 953 bp mRNA linear EST 02-JUL-2004
DEFINITION EST844419 Sequencing ESTs from 10b101ly pine embryos Pinus taeda

ACCESSION CDNA clone P1AM182 5' end, mRNA sequence.
 VERSION COA14034
 KEYWORDS COA14034.1 GI:49630282
 SOURCE EST.
 ORGANISM Pinus taeda (loblolly pine)
 Pinus taeda
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
 1 (bases 1 to 953)
 AUTHORS Buell, C.R., Heiao, J., and Cairney, J.
 TITLE Sequencing of ESTs from loblolly pine embryonic libraries
 JOURNAL Unpublished (2004)
 COMMENT Contact: C. Robin Buell
 Plant Genomics Group
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: rbuell@tigr.org
 This clone is available through TIGR. Please contact pine@tigr.org
 for further information
 Seq primer: APT TAG GTG ACA CTA TAG.
 Location/Qualifiers

FEATURES

SOURCE

1..953
 /organism="Pinus taeda"
 /mol_type="mRNA"
 /cultivar="7-56 mother tree, open-pollinated tree from,
 Lyons, Georgia, USA"
 /db_xref="taxon:3352"
 /clone="P1AM182"
 /lab_host="E.coli DH10B-TonA"
 /clone_1b="Sequencing ESTs from loblolly pine embryos"
 /note="Organ: Zygotic Embryo and Megagametophyte, Somatic
 Embryo; Vector: pCMV-SPORT 6.1; Site_1: NotI; Site_2:
 EcoRV; tissue: Whole megagametophytes isolated from pine
 seeds, whole embryos excised from these megagametophytes,
 whole somatic embryos and suspensor tissue from tissue
 culture, isolated from cell line A12. Pooled RNA from
 zygotic embryos, megagametophytes, and somatic embryos was
 used for library construction. Pine cones were harvested
 weekly from open-pollinated 7-56 mother trees, collections
 occurred from 7/01/02 until 10/15/02. Whole
 megagametophytes were first isolated from pine seeds, and
 whole embryos excised from these megagametophytes each was
 flash frozen. Embryo development was assessed using the
 system of Pullman et al (Pullman GS, Johnson S, Peter G,
 Cairney J, Xu N. 2003. Loblolly pine somatic
 embryogenesis: development of a maturation medium and
 resulting embryo quality. Plant Cell Reports 21:747-758
 (http://link.springer.de/link/service/journals/00299/conte
 nts/03/00586/). For photographs see Ciavatta et al 2001.
 (Ciavatta VT, Morillon R, Pullman GS, Chrispeels M,
 Cairney J. 2001. An aquaglyceroporin is abundantly
 expressed early in the development of the suspensor and
 the embryo proper of loblolly pine (Pinus taeda L.). Plant
 Physiol. 127: 1556-1567
 (http://www.plantphysiol.org/cgi/content/full/127/4/1556))

ORIGIN

Alignment Scores:
 Pred. NO.: 3.59e-162 Length: 953
 Score: 1422.00 Matches: 273
 Percent Similarity: 94.62% Conservative: 26
 Best Local Similarity: 86.35% Mismatches: 17
 Query Match: 54.05% Indels: 1
 DB: 7 Gaps: 0

US-10-718-952-12 (1-510) x COA14034 (1-953)

QY 115 SerPheGlnGlyGluGluIleTyrAlaProPheIysSerLeuLeuProMetValAsnPro 134
 Db 7 TCCTTTACGGCGAAGAGATTCATCTCTTCAAGAGCCTTCTCCATGCTGATCCT 66
 QY 135 AspAspIleValPheGlyGlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetVal 154

Db 67 AATGAATCGCTCTTTGGTGGATGGACATTAACAACATGAATCTAGCCGATCCATGGCC 126
 QY 155 ArgAlaIysValPheAspIleAspLeuGlnIleGlnLeuAArgProTyrMetGluSerMet 174
 Db 127 CGAGCCAAAGGTTCTTGACATTGACCTACAGAGCGCTCGCGCTTACATGCAGGCAATG 186
 QY 175 ValProLeuProGlyIleTyrAspProAspPheIleAlaIAsnGlnGluIuArgAla 194
 Db 187 ACTCTCTTCTCGTATATATGACCCGATTTCTATAGCGGCACCAACAGGATGAGGGCA 246
 QY 195 AsnAsnValIleIysGlyThrIysGlnGluGlnValGlnGlnIleIleIysAspIleIys 214
 Db 247 AATATGTCATCAAAAGCTCCAGAAAGCAACGCAAAAGATTATCCAAATATCAGCA 306
 QY 215 AlaPheIysGluAlaThrIlysValAspIysValValIleTyrThrAlaAsnThrGlu 234
 Db 307 GATTTTAATCGAAMAACAAGGTGACAAAGTAGCTGACTGTGACTGATTAATCCGAA 366
 QY 235 ArgTyrSerAsnLeuValIValGlyLeuAsnAspThrMetGluAsnLeuAlaIAlaVal 254
 Db 367 CGATACAGCGAGTATATGAGGTTTAATGACAGAAAGAAACCTTACTGCTCTCTG 426
 QY 255 AspArgAsnGluAlaGluIleSerProSerThrIleuTyrAlaIleAlaCySerValMetGlu 274
 Db 427 GAAGAATGATGATCGGAATCTCGGCATCGACTCTCTAGCGGTTGGCTTGCAAGAG 486
 QY 275 AsnValProPheIleAsnGlySerProGlnAsnThrPheValProGlyLeuIleAspLeu 294
 Db 487 AACGTTCCCTTCATTAACGGCAGTCTCTCAAAACCTTCGTTCCAGATTGATCATTTTA 546
 QY 295 AlaIleAlaAsnThrLeuIleGlyGlyAspAspPheIysSerGlyIleThrIysMet 314
 Db 547 GCAATTACAAAGAACAGATATATCGACAGAAATGATCTTCAAGAGCGGCGACAAAGATG 606
 QY 315 LysSerValIleuValAspPheLeuValGlyIleIleIysProThrSerIleValSer 334
 Db 607 AAGTCGTACTCGTGATTTCTTACTGCTCGCTCTTAAGCCCACTTCATTGTGAGC 666
 QY 335 TyrAsnHisLeuGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSer 354
 Db 667 TACATCATCTCGGTAAACAAGCGGATGATGTGACGCTCCCAACATTCGCGTCA 726
 QY 355 LysGluIleSerIysSerAsnValValAspAspMetValAsnSerAsnAlaIleuIys 374
 Db 727 AAGGAAATTTCAAGAGCAATGCTGTCCACACATGTTGCCAGAACCCATTCCTTAC 786
 QY 375 GluProGlyGluHisProAspHisValValIleIysTyrValProTyrValGlyAsp 394
 Db 787 GAACCAAGGGAACCCAGACCACTGCTGCTCATCAAGTATGTACTTACCTTGAGAGC 846
 QY 395 SerIysArgAlaMetAspGluTyrThrSerGluIlePheMetGlyIlyAsnAsnThrIle 414
 Db 847 AGCAAGAGAGCCATGAGACGATGACACGTCGAGATATTATGGAGGAAA-AAACCCCTT 905
 QY 415 ValLeuHisAsnThrCysGluAspSerLeuLeuAlaIAspIleIle 430
 Db 906 GTAATCCACAATATCATGTGAAGATTGCTTGTGAGCTCCTGTGATC 953

RESULT 12
 LOCUS CO113870 898 bp mRNA linear EST 16-JUN-2004
 DEFINITION GR_EB014121.r GR_Eb Gossypium raimondii cDNA clone GR_EB014121
 3', mRNA sequence.
 ACCESSION CO113870
 VERSION CO113870.1 GI:48812557
 KEYWORDS EST.
 SOURCE Gossypium raimondii;
 ORGANISM Gossypium raimondii;
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 1 (bases 1 to 898)

AUTHORS

Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and
Wing, R.A.

TITLE

Global assembly of Cotton ESTs

JOURNAL

Unpublished (2004)

COMMENT

Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 014 row: 1 column: 21.

FEATURES

source

Location/Qualifiers

1..898

/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_EB014121"

/tissue_type="floral"

/dev_stage="3 to +3 DPA"

/lab_host="DH10B"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:

EcoRV; Library made by Invitrogen with RNA supplied by

Wendle lab. Directional cloned into NotI-EV. Colonies

plated/picked by AgI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:

Pred. No.: 1.99e-158 Length: 898
Score: 1391.00 Matches: 263
Percent Similarity: 95.65% Conservative: 23
Best Local Similarity: 87.96% Mismatches: 13
Query Match: 52.87% Indels: 0
DB: Gaps: 0

US-10-718-952-12 (1-510) x COI13870 (1-898)

49 ValaentYrGlnPhelYsThraNthrhHsValProLysLeuGlyValMetLeuValGly 68
Db 1 GTCAATATGAAATTCAAGACTGATACCCATGCCCTTAATTGGGGTGATGCTTGGA 60
Qy 69 TTPGlyGlyAsnAsnGlySerThrLeuThrglyValIleAlaAsnArgGluAspIle 88
Db 61 TGGGAGAGAAACAATGGTTCAACCTCACCGGTGTTATAGCTAACAAAGAGGTATC 120
Qy 89 SerTPaathrLysAspLysIleGlnGlnAlaAsnYrPhelGlySerLeuThrGlnAla 108
Db 121 TCTTGGGTACTAAAGACAGAGTCAACAGCGCTAATTGCTTCAATGACTCAACGA 180
Qy 109 SerAlaIleArgValGlySerPheGlnGlyGluGluIleYrAlaProPheLysSerLeu 128
Db 181 TCACGATCCCAATGGCTTTCACAAATGAGAGAGATTATGCTCCATTTAAGGCTCTT 240
Qy 129 LeuProMetValAsnProAspAspIleValPheGlyGlyTPAspIleSerAsnMetAsn 148
Db 241 CTTCTTATGTTGTAACCAATGATATGTTGTTGAGAGTGGACATTAGTACATGAAC 300
Qy 149 LeuAlaAspAlaMetAlaArgAlaLysValPheAspIleAspLeuGlnLysGlnLeuArg 168
Db 301 CTAGCTGATGCATGGCTTAGGGCCAAAGTTTTCGATGCATGCATCTCCAAAGCACTGAGA 360
Qy 169 ProYrMetGluSerMetValProLeuProGlyIleYrAspProAspPheIleAlaAla 188
Db 361 CCTTCAATGGAATCCATGCTCCACTCCCTGGAATCTAGATCTGATTTCTTCTGCT 420
Qy 189 AsnGlnGluGluArgAlaAsnAsnValIleLysGlyThrLysGlnGluGlnValGlnGln 208
Db 421 AACCAAGGTGAACGTGCCAATATGTCATCAAGGGAGCAACAAGAAAGCAAGTTCAAGCAG 480
Qy 209 IleIleLysAspIleLysAlaPheLysGluAlaThrLysValAspLysValValLeu 228

Db 481 GTCAATCAAGCATCAAGAGCTTCAAGAGAAAAACAAGGTGACCAAGCTTGTACTC 540
Qy 229 TTPThAlaAsnThrGluArgYrSerAsnLeuValGlyLeuAsnAspThrMetGlu 248
Db 541 TGAATGCAAAACACTGAGAGAGTACAGCAATGTCATCGTGGGCTTAATGACACCGTGA 600
Qy 249 AsnLeuAlaAlaValAspArgAsnGluAlaGluIleSerProSerThrLeuYrAla 268
Db 601 AGCCTTAGGCTTCTTGGAGAGAAATGAATCAAGATTCTCTTCACCTTGATAGCT 660
Qy 269 IleAlaCysValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheVal 288
Db 661 ATGCTTGTTGTTCTTGAATATGTTCTTCTTCAATGACACCAACAAACACTTGTCT 720
Qy 289 ProGlyLeuIleAspLeuAlaIleAlaArgAsnThrLeuIleGlyYrAspAspPheLys 308
Db 721 CCAGGGTTGATTGATTGGCTATTCAAGAGAACTGCTGATTGAGAGATGACTTCAAG 780
Qy 309 SerGlyGlnThrLysMetLysSerValLeuValAspPheLeuValGlyAlaGlyTleLys 328
Db 781 AGTGGCCAGACCAAGATGANAATGTCCTGCTGATATTCCTTGTGGGGCTGGGATCAAG 840
Qy 329 ProThrSerIleValSerYrAsnHsLeuGlyAsnAsnAspGlyMetAsnLeuSer 347
Db 841 CCAACATGATGATGAGTTTACACCATCTGGGAAATATGATGCAATGATCTGCA 897

RESULT 13
CO085839
LOCUS
DEFINITION GR_Ea03E19.f GR_Ea Gossypium raimondii cDNA clone GR_Ea03E19.5',
mRNA sequence.

ACCESSION CO085839
VERSION CO085839.1 GI:48764473
KEYWORDS EST.

SOURCE
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Malvales; Malvaceae; Malvaceae; Gossypium.

REFERENCE
1 (bases 1 to 865)
Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and
Wing, R.A.

TITLE
JOURNAL
COMMENT Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 03 row: E column: 19.

FEATURES
source
1..865
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea03E19"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_id="GR_Ea"

ORIGIN
Alignment Scores:
Pred. No.: 1.69e-156 Length: 865
Score: 1375.00 Matches: 264
Percent Similarity: 97.22% Conservative: 16

US-10-718-952-12 (1-510) x COI13870 (1-898)

QY 303 Glycylasparaphenylserglycylthrlysmetylservalleuvalasphenleu 322
DB 491 GGTGTGACGACTTCAAGAGTGTGACAGACCAAGAAATGCTGTCGATTCCTT 550
QY 323 Valglyalaglyllelyserprothrseryllevalsertryasnhielseuglyasnansp 342
DB 551 GTTGTGCTGGAATTAAGCCCACTCAATCGTAAGCTTACACCACTTGGAAACAAGAT 610
QY 343 GlymetasleuSERalaproglinthrphearserlysguileserlyserasnval 362
DB 611 GGCATGAACCTGTCTGCCCTCAACATTCAGGTCCAAAGAGATCTCCAAAGCAACG 670
QY 363 Valasparpmetvalasmsersasnalaileuerygluproglygluhisproasphs 382
DB 671 GTGGATGACATGTCTCGACGATGCCATCCCTCTGTAGCCCGGAGACATCCCATCAT 730
QY 383 ValvalalallelyserValprotyrValglyaspserysargAlmetAspGluTyr 402
DB 731 GTCTGTTCATCAATATATGTCCGTAAGTGGAGACAGACAGGAGCTATGACAGATAC 790
QY 403 ThrseryllepheMetGlyglylysaanthrillevalleuhiAsnthrCysGluasp 422
DB 791 ACCTAGAGATCTTATGGGCGGCAAGAACCATCGTGTGCAACAACCTGTGAAGAC 850
QY 423 SerleuLeuAlaAlaProIlelleuAspleu 433
DB 851 TCGCTCTCGCCGCACTATCATCTTGATCTG 883
RESULT 15
CV290142 891 bp mRNA linear EST 23-SEP-2004
DEFINITION aoF01-10ms3-d10 AoF01 Asparagus officinalis cDNA clone
ACCESSION CV290142
VERSION CV290142.1 GI:52575144
KEYWORDS EST
SOURCE Asparagus officinalis (garden asparagus)
ORGANISM Asparagus officinalis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
Asparagaceae; Asparagus.
1 (bases 1 to 891)
dePamphilis,C., Carlson,J., Ma,H., Soltis,D., Soltis,P.,
Openheimer,D., Frohlich,M., Doyle,J., Tanksley,S., Webb,M.,
Leebens-Mack,J., Landherr,L., Ilut,D., and Wall,K.
Generation of ESTs from early male inflorescences of Asparagus
officinalis
JOURNAL Unpublished (2004)
COMMENT Contact: Claude dePamphilis or James Leebens-Mack
Muller Laboratory
Penn State University
208 Muller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 863 6413
Fax: 814 865 9131
Email: cwj3@psu.edu or jhl10@psu.edu
The sequence provided is trimmed of vector and low quality regions.
Full sequence and original trace file are available from the Plant
Genome Network website (http://pgn.cornell.edu)
Plate: aoF01-10ms3 row: d column: 10
Seq primer: M13F.
FEATURES
source
Location/Qualifiers
1..891
/organism="Asparagus officinalis"
/mol_type="mRNA"
/db_xref="taxon:4686"
/clone="aoF01-10ms3-d10"
/tissue_type="male inflorescences"
/lab_host="SOLR"
/clone_1b="AoF01"
/note="Vector: pBluescript SK (+/-); Site 1: EcoRI;
Site 2: XhoI; This is a directionally cloned,
non-normalized library. This library has been generated by

the Floral Genome Project (FGP). The Floral Genome Project
is funded by NSF's Plant Genome Research Program
(DBI-0115684). More information about the project can be
obtained at <http://fgp.bio.psu.edu>

ORIGIN
Alignment Scores:
Pred. No.: 2,34e-156 Length: 891
Score: 1374.00 Matches: 268
Percent Similarity: 97.58% Conservative: 14
Best Local Similarity: 92.73% Mismatches: 7
Query Match: 52.22% Indels: 1
DB: 7 Gaps: 0
US-10-718-952-12 (1-510) x CV290142 (1-891)
QY 222 ValAspLyValValleuTTrThAlaAnThrGUAArgTYrSerAsnLeuVal 241
DB 4 GTAGCAAGATGATGGTCTTTGGATCGCAATTCGAGAGGTATAGCAATGTGATTC 63
QY 242 GlyleuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAla 261
DB 64 GGTCTCATGACACATGAGAGAACCTTGCTGCGCTGTGACAAAGAACAGACAGATA 123
QY 262 SerProserThrleuTYrAlaIleAlaCysValMetGluAsnValProPheIleasnGly 281
DB 124 TCCCATCAACCCCTATGATGCTTGCTGCTGATGGAGAACCTTCCATCATCAACGA 183
QY 282 SerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThrleu 301
DB 184 AGCCCAAGAACACCTTCTGCTGCTGTTGATGATTTGGCTATTCAAGAAAGATCTTA 243
QY 302 IleGlyGlyAspAspPheLySerGlyGlnThrlysmetylservalleuValAspPhe 321
DB 244 ATGGAGAGATGACTTTAAGATGACAGACCAAGATGAAGTCGTGTGATTC 303
QY 322 LeuValGlyAlaGlyllelyserProthrseryllevalsertryasnhielseuglyasn 341
DB 304 CTGTGGGGGTGGAAATCAACCACTTCATTTGTAGCTTCAACCACTTGGGAACAT 363
QY 342 AspGlyMetAsnleuSERalaproglinthrphearserlysguileserlyserasn 361
DB 364 GATGGCATGAATCTCTCGCAACCAACCTTCCGATCAAGAGATTCGAAGACAT 423
QY 362 ValValAspAspMetValasmsersasnalaileuerygluproglygluhisproasp 381
DB 424 GTGGTCATGACATAGGTTTCAAGCAATGGCATCTGTATGAGCTGGGAGACACCTGAT 483
QY 382 HisValValVallelyserTYrValProTYrValGlyAspserlyBAgAlaMetAspGlu 401
DB 484 CATGCATTTGTTATCAAGTATGTCCGTATGTGTGTATGACCAAGAGGCTATGGATGAG 543
QY 402 TyrThrserGlyllePheMetGlyglylysaanthrillevalleuhiAsnthrCysGlu 421
DB 544 TACACATTTGATTAATTCATGAGAGTCCAGCCACCATTTGTCTGCACAAACCTGTAG 603
QY 422 AsperleuLeuAlaAlaProIlelleuAspleuValleuAlaGluLeuSerThr 441
DB 604 GATTCGCTCTTGGCTGGCGGCGATCATCTGACTTGGTTGATTGGCTGAGACTTAGACC 663
QY 442 ArgIleGlnPheLyAlaGluAsnGluGlylyspPheHiserPheHisProValAlaThr 461
DB 664 AGGATTCAGCTCAAGGCTGAGAGAGAGGCAAGTCCACTCTTCATCCAGTGGCTAAC 723
QY 462 IleleuSerTYrleuThrlyValaProleuValProProGlyTYrProValValAsnAla 481
DB 724 ATCTTGATTAACCTCACCAAGGCTCCCTGTTGTCCACAGGACACCAAGTGTGAATCA 783
QY 482 LeuSerlysglnAGAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAlaPro 501
DB 784 CTCTCAAGCAAGAGGAGATCTGAGAAACATCTCGAGGCTTGCAATTGGCTTGCTCT 843
QY 502 GluAsnAsnMetIleleuGlyTyrlys 510

DB 844 GAGAAC-ATGATCCTGAGTACAA 869

Search completed: June 7, 2005, 22:49:35
Job time : 3751.5 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: June 7, 2005, 17:22:56 ; Search time 219 Seconds
(without alignments)
3810.507 Million cell updates/sec

Title: US-10-718-952-12
Perfect score: 2631
Sequence: 1 MFIEFKYSPVVKYTTETI.....NIMRACVGLAPENNMLEYK 510

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/US10718952/rnat.06062005.173401.12868/app_query.fasta_1.1294
-DB=Issued_Patents_NA -QFMT=fastp -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORMaxext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10718952@cgn1_1.141@rnat.06062005.173401.12868 -NCPU=6 -ICPU=3
-NO MMAP -LANG=EDUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONCLDLOC
-BEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfillseq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2346	89.2	1931	3	US-09-118-442-10
2	2346	89.2	1931	3	US-09-118-442-10
3	2346	89.2	1931	3	US-09-118-442-10
4	1823	69.3	3546	3	US-09-118-442-15
5	1823	69.3	3546	3	US-09-118-442-15
6	1823	69.3	3546	3	US-09-118-442-15
7	1822	69.3	3546	3	US-09-118-442-14
8	1389.5	52.8	1578	4	US-09-734-2378-72
9	1389.5	52.8	1578	4	US-09-734-2378-72
10	1384.5	46.7	77626	4	US-09-248-796A-1311
11	1229.5	41.9	1231	3	US-09-949-016-12608
12	1103.5				Sequence 12608, A
					Sequence 34, Appl

c	13	529	20.1	34316	4	US-09-902-540-1257	Sequence 1257, Ap	
	14	528	20.1	1335	4	US-09-902-540-5256	Sequence 5256, Ap	
	15	160	6.1	294	4	US-09-313-294A-4684	Sequence 4684, Ap	
	16	142	5.4	42325	4	US-08-311-731A-131	Sequence 131, App	
	c	17	128.5	4.9	4403765	3	US-09-103-840A-2	Sequence 2, Appl
	c	18	127.5	4.8	4411529	3	US-09-103-840A-1	Sequence 1, Appl
	c	19	119.5	4.5	1359	4	US-09-248-796A-6542	Sequence 6542, Ap
	c	20	118.5	4.5	8930	4	US-09-077-098A-1	Sequence 1, Appl
	c	21	118	4.5	7571	3	US-08-961-527-2	Sequence 2, Appl
	c	22	117.5	4.5	274	3	US-09-118-442-21	Sequence 21, Appl
		23	117.5	4.5	274	3	US-09-677-064-21	Sequence 21, Appl
		24	116	4.4	2190	4	US-09-107-433-2104	Sequence 2104, Ap
		25	116	4.4	2196	4	US-09-583-110-2059	Sequence 2059, Ap
		26	115	4.4	2529	4	US-09-489-039A-4399	Sequence 4399, Ap
		27	114	4.3	2193	1	US-09-014-897-3	Sequence 1, Appl
		28	114	4.3	2193	1	US-09-014-897-3	Sequence 1, Appl
		29	114	4.3	2193	1	US-08-731-716-3	Sequence 1, Appl
		30	114	4.3	2193	1	US-08-731-716-3	Sequence 3, Appl
		31	113.5	4.3	2337	4	US-09-489-039A-1284	Sequence 1284, Ap
		32	112	4.3	3421	4	US-08-956-171E-337	Sequence 337, App
		33	112	4.3	3421	4	US-08-781-966A-337	Sequence 337, App
		34	110.5	4.2	2244	4	US-09-583-110-405	Sequence 405, App
		35	110.5	4.2	2256	4	US-09-107-433-772	Sequence 772, App
		36	110.5	4.2	5892	4	US-09-583-110-2582	Sequence 2582, App
		37	110.5	4.2	5904	4	US-09-107-433-2280	Sequence 2280, App
	c	38	110.5	4.2	640681	4	US-09-790-988-1	Sequence 1, Appl
	c	39	109.5	4.2	1230230	4	US-09-438-185A-1	Sequence 1, Appl
	c	40	109	4.1	8287	4	US-09-023-655-1441	Sequence 1441, Appl
		41	108.5	4.1	3729	4	US-09-107-532A-1587	Sequence 1587, App
		42	107	4.1	1052	3	US-08-961-527-244	Sequence 244, App
		43	106.5	4.0	2563	4	US-09-016-434-1076	Sequence 1076, App
		44	106.5	4.0	2563	4	US-09-023-655-894	Sequence 894, App
		45	106.5	4.0	2601	3	US-08-569-749-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-118-442-10
; Sequence 10, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-10
Alignment Scores:
Pred. No.: 4.26e-285
Score: 2346.00
Percent Similarity: 97.33%
Best Local Similarity: 87.84%
Query Match: 89.17%
DB: 3
Length: 1931
Matches: 448
Conservative: 28
Mismatch: 34
Indels: 0
Gaps: 0

```
US-10-718-952-12 (1-510) x US-09-118-442-10 (1-1931)
QY 1 MetPheIIeGLuaAspMetValGluSerProAsnValIleTyrThrGluThrGluIle 20
Db 99 AGGTTTCATGAGAGCTTCGCGTCGAGAGCCCCCGGATGAGCGCCGATGAGATC 158
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 159 GAGTGGAGTACCGGATGACAGACAGAGAGCTGTATCAGAGAGGCGAAGAGCGCGCTCA 218
QY 41 GlnTyrIleValIleAspProIleSerValAsnTyrGlnPheIleThrAsnThrHisValPro 60
Db 219 CCTGGGTGCTCGCGCCCAAGTCCGTCAGATCAATCTCCGAGACCAAGACCCCGCTCCC 278
QY 61 LysLeuGluValMetLeuValGlyTyrPheGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db 279 AAGCTCGGGGTATGCTTGTGGGGTGGGAGGCAACAGGGGTCCACGCTGACGGCTGGG 338
QY 81 ValIleAlaAsnArgGluAspIleSerTyrAlaThrLysAspLysIleGlnGlnAlaAsn 100
Db 339 GTCATTTGCCAAGAGGAGGAGATCTCATGGCGGACCAAGAGCAAGGTGACGACCAAC 398
QY 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
Db 399 TACTACGGCTCCCTCCACCGCCTCCACATCAGAGTGGCAGCTCAACAGGGAGAGAG 458
QY 121 IleTyrAlaProPheIleSerLeuLeuProMetValAsnProAspIleValPheGly 140
Db 459 ACTTAAGCCCGCTCAAGAGCCCTCTCCATAGTAAGAACCCAGACCACTTGTGTGGGA 518
QY 141 GlyTyrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
Db 519 GGCTGGGACATTATGACACATGAACCTGGCGCATCTCATACAGGCGCAAGGTGCTGAT 578
QY 161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProGlyIle 180
Db 579 ATTGACCTGACAGAGAGCTCAGGCGCTCATGAGATGATGATGATGATGATGATGATG 638
QY 181 TyrAspProAspPheIleAlaAlaAsnGlnGluGluArgAlaAsnAsnValIleLysGly 200
Db 639 TATGATCCCGAGCTTCACTCGCGCTAACAGGGCTCTCGGCGCAAGCTGTCTCATCAAGGGC 698
QY 201 ThrLysGlnGluValGlnGlnIleIleLysAspIleLysAlaPheLysGlnAlaThr 220
Db 699 ACCAAGAAAGAAAGAGGTGAGAGCATCATCAAGAGATATAGGAGATTTAAGAGAAAGAAC 758
QY 221 LysValAspLysValValIleTyrThrAlaAsnThrGluArgTyrSerAsnLeuVal 240
Db 759 AAAGTGCACAAAGATGTTGTTGTGTGACTGCAAAACACTGAAGAGTATGCAATGTGTGC 818
QY 241 ValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaValAspArgAsnGluAlaGlu 260
Db 819 GCTGGCTCAAGACAGATGAGATCTACCTGAGCATCTGTGAGACAAAGAACAGCGGAG 878
QY 261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
Db 879 GATACCACTCAACACTATATGCTGCTGTGTATGAAAGGGGTGCCGTTCAAT 938
QY 281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
Db 939 GGGAGGCCCCAGAACACTTTGTGCTGGCTGATGATCTTGCTATATAAAAAACAACCTGC 998
QY 301 LeuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
Db 999 TTGATTTGGTGTGACACACTTCAAGAGTGAAGAGACCAAAATGAAATCTCTTTGGTGCAT 1058
QY 321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnHisIleGluLysn 340
Db 1059 TTCTCTTTGTTGCTGGAATTAAGCCCACTCAATGTAGCTACCAACCACTTGGGAAAC 1118
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
Db 1118
```

```
Db 1119 AACGATGCGATGAACCTGTCTGCCCTTCAACATTCAGGTCCAAAGAGATCTCCAAAGC 1178
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
Db 1179 AAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1238
QY 381 AspHisValValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
Db 1239 GATCATGTGTTGTATCATCAAGATATGTCGTAACGTGGGAGACAGACAAAGAGGCTATGAC 1298
QY 401 GlnTyrThrSerGluIlePheMetGlyGlyLysAsnThrIleValLeuHisAsnThrLys 420
Db 1299 GAGTACACCTCAGAGATGTTATGAGGCGGAGAAACACATCTGTGCTGCAACACCTGT 1358
QY 421 GluAspSerLeuLeuAlaProIleIleLeuAspLeuValLeuLeuAlaGluLeuSer 440
Db 1359 GAGGACTGCTCTCGCGCCACCTTATCTTGTATCTGTGTGCTTGTGTGCTGAGCTCAGC 1418
QY 441 ThrArgIleGlnPheLysAlaGluAsnGluLysPheHisSerPheHisProValAla 460
Db 1419 ACCAGGATCCAGCTGAAGCTGAGGAGAGAGCAAAATTCACCTCTCCACCGGTGACC 1478
QY 461 ThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThrProValValAsn 480
Db 1479 ACCATCTTGAAGTACTTCCACCAAGGACCCCTGTCTCCCTGGCAGACCGGTGTGAAC 1538
QY 481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
Db 1539 GCTGTGGCAAGCAGAGGCGCATGTGTGAGAAACATCATGAGAGGCTGCTGTGGCTGGCC 1598
QY 501 ProGluAsnAsnMetIleLeuGluTyrLys 510
Db 1599 CCAAGAAACAAATGATCTTGGAGTACAG 1628
```

RESULT 2

```
US-09-677-064-10
; Sequence 10, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phytochemical Metabolism in
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-10
```

Alignment Scores:

```
Pred. No.: 4.26e-285 Length: 1931
Score: 2346.00 Matches: 448
Percent Similarity: 93.33% Conservative: 28
Best Local Similarity: 87.84% Mismatches: 34
Query Match: 89.17% Indels: 0
DB: 3 Gaps: 0
```

US-10-718-952-12 (1-510) x US-09-677-064-10 (1-1931)

QY		1	MeeThelIegLuanPheLyValGlnSerProAsnValLysTYrThrGlutHglulle	20
Db		99	ATgTTCATCGAGAGCTTCGGGTGCAGAGCCCAACGTCGGTAAGGCCCATGAAGATC	158
QY		21	GlnSerValTYrAsnTYrGLutThrThrgLluleuValHisGLuAsnArgAsnGLYrTYrX	40
Db		159	GAGTCGGAGGTACCGETTAGCACACGACGGAGCTGTATACAGAGGGCAAGACGGCGCTCA	218
QY		41	GlnTrpLleValLysProLysSerValAsnTYrGlnPheLysTHrAsnTHiEValPro	60
Db		219	CGCTGGGTGGTCCGCCCAAGTCCGTCAAGTAACTTCCGAGCACAGACCGCGCTCCC	278
QY		61	LysLeuGLyValMetLeuValGLYrTrpLYGlyAsnAnGLYserThrLeuThrGLYgly	80
Db		279	AAGCTCGGGGGATGCTTGTTGGGGGGAGACAACGGGTCCACCGCTAAGCTGCTGG	338
QY		81	ValILealaaAnaArgLuAspLLeserTrpaLaThrLysAspLysILeGLInlaAlasn	100
Db		339	GTACTTGCCAACAGGAGGGAGATCTCATGGCGACCAAGACACAGTCTCACCAACCA	398
QY		101	TyrPheGLYserLeuThrGlnAlaseralalleargValGLYserPheGLngLYgluL	120
Db		399	TACTACGGCTCCCTCACCCACGGCTCCACATCAGAGTCGGACAGCTAACGGGGAGAG	458
QY		121	ILetTYrAlaProPhelyYSerLeuLeuProMetValAsnProAspAPlleValPheLY	140
Db		459	ATCTATGGCGCGTTCMAAGAGCTCTCTCCATAGGAACCAAGACGATTTGTTCGA	518
QY		141	GLYrTPAspILeserAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheaSP	160
Db		519	GGCTGGGACATTAGCAACATAACCTGGCCCATCTCATGACCAAGGGCCAAAGTGTGAT	578
QY		161	ILeaspleuglnLYsglnLeuArgProTYrMetGlnSerMetValProLeuProGLYlle	180
Db		579	ATTGACCTCGAGAAGCAAGCTCAGGGCCCTAATAGAGTCATCAGTGCCTCCCGTATC	638
QY		181	TYrAspProAspPheIlleAlaAlaAsnGLngluAlaAlaAsnValILleYsgLY	200
Db		639	TATATATCGGACTTCATCGCGGCTTAAACAGGGCTCTCCGGCCAACGTGTCATCAAGGCG	698
QY		201	ThrlYsglnglnLgnValGlnGlnIlelleLysAspILleYsAlaPheLYsglnAlaThr	220
Db		699	ACCAAGAAAGAACAGGTGAGCAAGATCAACAGGATATCAAGGAGTTTAAAGAAAGAAC	758
QY		221	LysValAspLysValValleuTrpThraAlaAsnTrpGLuArgTYrSerAsnLeuVal	240
Db		759	AAAATGGACAAAGATATGTTGTGTGTGGACTGCCAACAACCTGAAGAATATAGCAATGTGCG	818
QY		241	ValGLYLeuAsnAspThrMetGlnAsnLeuLeuAlaValAspArgAsnGLuaGLu	260
Db		819	GCTGTGTCMAACGACACAGATGAGATATACGTGCATCTGTGGACAAAGAACGGAG	878
QY		261	ILeserProserThrLeuTYrAlaIlleAlaCySValMetGluAsnValProPheIleSn	280
Db		879	GTATCCATCAATCAACATATATGCACTTGCCTGTGCATGAGAAAGGGTCCCTCATCATAT	938
QY		281	GLYserProGlnAsnThrPheValProGLYLeuIleAspLeuAlaIlleAlaArgAsnThr	300
Db		939	GGAGGCCCCCAAGAACCTTTGTGCTGGCTGGCGATGATCTTGCTATAAAAACAACCTGC	998
QY		301	LeuIleGLYglYaAspAspPheLysSerGLYInthrLysMetLysSerValleuValAsp	320
Db		999	TTGATTTGGTGTGACGACTTCAAGAGTGGACAGACCAAGATGAATCTGTCTTGGTGGAT	1058
QY		321	PheLeuValGLYAlaGLYILleYsVerProThrsertlleValserTYrAnHisleuGLYsn	340
Db		1059	TTCTCTTGTGTGGTGGAAATAAAGCCCACTCAATCGTAGACTCAACACCTTGGGAAC	1118
QY		341	AsnAspGLYmetAsnLeuSerAlaProGlnThrPheArgSerLYsglnILEserLysSer	360
Db		1119	AAACATGGCATGAACCTGTCTGCCCTCAACACATTCAAGTCCAGAGCATCTCCAAAGCC	1178
QY		361	AsnValValAspAspMetValAsnSerAnalaileuTYrGLuProGLYgluHisPro	380

Db	1179	AACTGTGTGATGACATGTTCTTCGAGCAATGCAATCCTCTATGAGCCCGGAGACATCC	122
Qy	381	AspHisValValValIleValSerValProValValGlyAspSerValArgAlaMetAsp	400
Db	1239	GATCATGTGCTTGTATCATCAATGATGTGCGTACGTGGAGACACCAAGAGGGCTATGAGAC	129
Qy	401	GluTyrThrSerGluIlePheMetGlyValTyrSerAsnThrIleValIleHisAsnThrCys	420
Db	1299	GAGTTACACCTCCAGAGAGTCTTCAATGGGCGGCAAGAACCATCTGGCTGCACACACCTGT	139
Qy	421	GluAspSerIleuLeuAlaIleProIleIleLeuAspLeuValIleuLeuAlaGluLeuSer	440
Db	1359	GAGAGCTCGCTCCTCGCCGCACTTATCATCTTGAATCTGGTGTCTTGGCTGAGCTAGC	144
Qy	441	ThrArgIleGlnPheValAlaGluAsnGluGlyIlePheHisSerPheHisProValAla	460
Db	1419	ACCAGAGTCCAGCTGAAAGCTGAGGAGGAGGACAAATTCACCTCCTTCCACCCGGTGGCC	147
Qy	461	ThrIleLeuSerTyrIleuThrTyrAlaProLeuValProProGlyThrProValValAsn	480
Db	1479	ACCATCTTGAATCTTACTTCCACCAAGGCACCCCTGGTTCCCTTGGCAACCCGGTGGTAC	153
Qy	481	AlaLeuSerTyrGluAlaArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla	500
Db	1539	GCTGTGGCCAGACAGAGGGCGATGCTGGAGAACATCATGAGGGCTGGGTGGCTGGCC	159
Qy	501	ProGluAsnAsnMetIleuGluTyrLys 510	
Db	1599	CCAGAGACACATGATCTTGGAGATACAG 1628	
RESULT 3			
US-09-727-628-1			
Sequence 1, Application US/09727628			
Patent No. 6791013			
GENERAL INFORMATION:			
APPLICANT: Armstrong, Katherine			
APPLICANT: Hey, Timothy D			
APPLICANT: Folkerts, Otto			
APPLICANT: Smith, Kelley A			
APPLICANT: Hopkins, Nicole L			
TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER			
FILE REFERENCE: 50597			
CURRENT APPLICATION NUMBER: US/09/727,628			
CURRENT FILING DATE: 2000-12-01			
PRIOR APPLICATION NUMBER: US 60/168,612			
PRIOR FILING DATE: 1999-12-02			
NUMBER OF SEQ ID NOS: 3			
SOFTWARE: PatentIn Ver. 2.0			
SEQ ID NO 1			
LENGTH: 1959			
TYPE: DNA			
ORGANISM: Zea mays			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (137)..(1699)			
US-09-727-628-1			
Alignment Scores:			
Pred. No: 5,97e-284			
Score: 2337.00			
Percent Similarity: 93.33%			
Best local Similarity: 87.84%			
Query Match: 86.83%			
Db:	4	Matches: 1959	
		Conservative: 448	
		Mismatches: 38	
		Indels: 0	
		Gaps: 0	
US-10-718-952-12 (1-510) x US-09-727-628-1 (1-1959)			
Qy	1	MetPheIleGluAsnPhelValGluSerProAsnValTyrThrGluThrGluIle 20	
Db	137	ATGTTATTCGAGAGCTTCGCGCTCCAGAGCCCCACATCGCGGATACGGCCGACGAGATC 196	
Qy	21	GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40	

```
Db      197 GAGTGGAGTACCGGTACGACGACGAGCGTGTGCAAGAGGCAAGGACCGGCTTATGAC 256
      41 GlnTPrIleValLeuPProLysSerValAsnThrGlnPheLysThrAsnThiStaPro 60
      257 GCGTGGGTCGTCCGCCCAAGTCGCTCAAGTCAACTTCCTGGACCAAGCCGCGGTCC 316
      61 LysLeuGlyValMetLeuValGlyTPrGlyAsnAsnGlySerThrLeuThrGlyGly 80
      317 AAGCTCGGGGTATCTTGTGGGTGGGAGGCAACAGGTTCAACCTGACCGGCTGG 376
      81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
      377 GTCATTCGCCAACAGAGGAGGATCTCATGGCGCAACAGAGCAAGAGTGCAGCAACCAAC 436
      101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySerPheGlnGlyGlu 120
      437 TACTACGGCTCCCTCAACCGGCTTCCACATCAAGAGTGCAGCTACCAACGGGAGAG 496
      121 IleTyrAlaProPheLysSerLeuPProMetValAsnProAspAspIleValPheGly 140
      437 ATATATGGCGCTTCAAGAGCTCTTACCCATGTGTGAACCCAGACGACTTGTGTGGA 556
      141 GlyTPrAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
      557 GGCTGGACATCAGACGATGAACCTGGCAGATGCATGACCAAGGACCAAGGTCGTGAC 616
      161 IleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProLysIle 180
      617 ATTTAGCTTGCAAAACAGCTCAGGCTTACATGAGTGCATGTGTGCTCCGCTGTC 676
      181 TyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnValIleLysGly 200
      677 TATGATCCGGAATTCATCGCCCTAACGAGGCTCTCGTGCACAAATCTCATCAAGGC 736
      201 ThrLysGlnGluGlnValGlnGlnIleIleLysAspIleLysAlaPheLysGluAlaThr 220
      737 ACCAAGAAAGAACAGGTGAGAGCATCATCAAAAGATATCAGAGGAGTTTAAAGAGAAAC 796
      221 LysValAspLysValValValLeuThrPThrAlaAsnThrGluArgTyrSerMetLeuVal 240
      797 AAAGTGGACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 856
      241 ValGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsnGluAlaGlu 260
      857 GGTGGTCTCAACGACATGAGATCTGTGGCATCTGTGACCAAGAACAGGCGGAG 916
      261 IleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValProPheIleAsn 280
      917 ATCTCGCCATCAACACTATATGCATGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 976
      281 GlySerProGlnAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArgAsnThr 300
      977 GGGAGCCGCCCAAGAACTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1036
      301 LeuIleGlyValAspAspPheLysSerGlyGlnThrLysMetLysSerValLeuValAsp 320
      1037 CTGATCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1096
      321 PheLeuValGlyAlaGlyIleLysProThrSerIleValSerTyrAsnIleLeuGlyAsn 340
      1097 TTTCTTTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1156
      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSerLysSer 360
      1157 AACGACGGGATGAACCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1216
      361 AsnValAlaAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyGluHisPro 380
      1217 AACGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1276
      381 AspHisValAlaValIleLysTyrValProTyrValGlyAspSerLysArgAlaMetAsp 400
```

```
Db      1277 GATCATGTGTTGTTCATCAAGTATGTCCGTATGTGGAGACAGTAAAGGCTTATGAC 1336
      401 GlnTyrThrSerGluIlePheMetGlyLysAsnThrIleValLeuHisAsnThrCys 420
      1337 GAGTACACATCAGAGATCTTATGAGCGGCAAGACACATCGTGTGTGTGTGTGTGTGTGT 1396
      421 GluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValIleLeuAlaGluLeuSer 440
      1397 GAGAGCTGCTCTTCGCGGACCGCATCATCTCGATCTGGGTCTGTGTGTGTGTGTGTGTGT 1456
      441 ThrArgIleGlnPheLysValGlyAsnGluLysPheHisSerPheHisProValAla 460
      1457 ACCAGGATCCAGTTAAACCTGAGGGAACGACAAAGTTCACCTCTTCACCCGGTGGCC 1516
      461 ThrIleLeuSerTyrLeuThrThyAlaProLeuValProProGlyThrProValAlaAsn 480
      1517 ACCATCTTACCTACTACCTACCAAGGACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1576
      481 AlaLeuSerLysGlnArgAlaMetLeuGluAsnIleMetArgAlaCysValGlyLeuAla 500
      1577 GCTCTTGCAAGCAGAGGCGATGCTGGAGAAATCATGAGGGCTTGCGTGTGTGTGTGTGTGT 1636
      501 ProGluAsnAsnMetIleLeuGluTyrLys 510
      1637 CCAGAGAACAAACATGATCTGTGAGTACAAAG 1666
```

RESULT 4

```
US-09-118-442-15
/ Sequence 15, Application US/09118442B
/ Patent No. 6197561
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Calt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ APPLICANT: Wang, Xun
/ APPLICANT: Bowen, Benjamin A.
/ TITLE OF INVENTION: Genes Controlling Phycate Metabolism in
/ FILE REFERENCE: 0706
/ CURRENT APPLICATION NUMBER: US/09/118,442B
/ EARLIER FILING DATE: 1998-07-17
/ EARLIER APPLICATION NUMBER: 60/055,446
/ EARLIER FILING DATE: 1997-08-11
/ EARLIER APPLICATION NUMBER: 60/055,526
/ EARLIER FILING DATE: 1997-08-08
/ EARLIER APPLICATION NUMBER: 60/053,944
/ EARLIER FILING DATE: 1997-07-28
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FASTSEQ for Windows Version 3.0
/ SEQ ID NO 15
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-118-442-15
```

Alignment Scores:

Pred. No.:	1,16e-218	Length:	3546
Score:	1823.00	Matches:	446
Percent Similarity:	40.86%	Conservative:	26
Best Local Similarity:	38.45%	Mismatches:	36
Query Match:	69.29%	Indels:	652
DB:	3	Gaps:	9

US-10-718-952-12 (1-510) x US-09-118-442-15 (1-3546)

```
QY      1 MetPheIleGluAsnPheLysValGluSerProAsnValLysTyrThrGluThrGluIle 20
      60 ATGTCATTCGAGAGCTTCGCGTCGAGAGCCGCCGTCGCGTACGCGCCACGAGATC 119
      21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
      120 GAGTGGAGTACCGGTACGACGACGAGAGTGTGTACAGAGGCAAGAGACGCGCTTCA 179
```

QY 41 GlnTPIleValIysProIysSerValIAsnTYrGlnPheIysThrAsnThrHiIstValPro 60
Db 180 CCGTGGGTGCTCGCCCAAGTCCGTCAAGTAACTTCGACCAAGACCGCCGTCCTCC 239
QY 61 LysLeuGlyI----- 63
Db 240 AAGCTCGGGTANGTAGAGATGAGAGGGCCCTAGCCCTCACTCTGTGAAACCTCCCTCTC 299
QY 63 ----- 63
Db 300 CCGTCTCAGTCAAAATCCCTCCGAGATCAACTGGTCGGCGTCCCTCTAAATCTTAA 359
QY 63 ----- 63
Db 360 TGAATAATCTTACTGCTTGGCTTGAAAGACGACCGTGTATTTGTGACAGCTACGACAC 419
QY 63 ----- 63
Db 420 ACTTGCCCATCGGGATGGGTCAAAATCAGCTGATTTGAAATTGCAATGGATGGCCCTT 479
QY 63 ----- 63
Db 480 TTCCATATTTGATCATCTTCGCTCCTAAGTCAATGATTACAGAAACGTCCTTCTCTC 539
QY 63 ----- 63
Db 540 TGAACCTTGTCTTAAAGCTTTTGTCTGTGACGTGAGCTGTATTCATTTGTTCATGTA 599
QY 63 ----- 63
Db 600 AGATCAAAATTCACAGCAGGAGCAGATGACGACAGACAGAACTCATTAACGTACAAATTGA 659
QY 63 ----- 63
Db 660 TACTAGATTACTGGCAAGTGTGATACGCGCAATCTGCATCTCGACCCCTTGTGT 719
QY 63 ----- 63
Db 720 TAATTCCTGTTCTTATGCAATGTTGCTTACGTGACGCTGTTGTGTATGATGATGACGC 779
QY 63 ----- 63
Db 780 TGTCAACCGCTTGTCTCTGTCTGACGAGATGATCCAACTTTTCTGTTCTGCTGTCAGC 839
QY 64 ValMetLeuValIglYTrpGlyIysAsnAsnGlySerThrLeuThrGlyValIleAla 83
Db 840 GTGATGCTTGTGGGGTGGGAGGCAACACGGGTCCAGCTGACGGCTGGGGTCAATTGCC 899
QY 84 AsnArgGlu----- 86
Db 900 ACGAGGAGTGAAGTAGTACTTAATTGTCTTATTTGCTTCCGTTGTTTCAGTTATTA 959
QY 87 ----- 92
Db 960 ATGGCCCTGACAGAACTGAATTTTGTGTGTGGCTGTTCAAGGGAGTCTCAAGCCGAC 1019
QY 93 LysAspLysIleGlnGlnIleAsnTYrPheGlySerLeuThrGlnIleAsnIleAlaIleArg 112
Db 1020 AAGGACAAAGTGCAGAACCACTTACACGGCTC-CTCAACCAAGCCCTCCACACATCAGA 1078
QY 113 ValIysSerPheGlnGlyIleGlnIleIleTYrAlaProPheIysSerLeuLeuProMet 131
Db 1079 GTCGGACGCTACACGCGGAGAGATCTATGCGCCGTTCAAGAGCTCTCTCCCATGGTA 1138
QY 131 ----- 131
Db 1139 ATCTATTAATAGCTTGAATTAATCTCTTTTACTGAAACCAACATATCAATAAAG 1198
QY 131 ----- 131
Db 1199 CATATTCGCTAAGGTCGTAGTTGATGTATTAAGTAACTGTCTTTCAGGCCAGTGGTC 1258
QY 131 ----- 131

Db 1259 TCAAGTAAACGGAATGTTAATCATTTGGGTGAAAAAACAAAGTTCTAATTTTGTGAAG 1318
QY 131 ----- 131
Db 1319 GAATGTTAACTTAGATTAATGAAAAAGGAGCAATTGTAAAGAAAGGTGTGAACAATC 1378
QY 131 ----- 131
Db 1379 GACTCGGTCTCCATGTTGTGATCTTACTTGCAAGTCAAAAAGTTCTGTGTAGCTCAA 1438
QY 131 ----- 131
Db 1439 AGGTTCCAGCATCTTTGGAATTACATCTCGTGCAGATTGACGATGCTGAATGCTGCA 1498
QY 131 ----- 131
Db 1499 GATTGCAGACTCGGTGTTGTGTATCTTCTTCATGACCAAGTGTGAACTGTTTCA 1558
QY 132 -ValAsnProAspAspIleValIlePheGlyIleAspIleSerAsnMetLeuAlaAs 151
Db 1559 GGTGAACCCAGACGACATTTGTGTGAGGCTGGACATTAAGCAACATGAACCTGGCCGA 1618
QY 151 PALAsnMetAlaArgAlaIysValIleAspIleAspLeuGlnIleAsnArgProTYrMe 171
Db 1619 CTCATGACCAAGGCCAAGTCTGATTTGACTGAGAAAGCAGCTCAAGGCCCTACAT 1678
QY 171 fGluSerMetValIProIeu-ProGlyIleTYrAspProAspPheIleAlaIleAsnGln 191
Db 1679 GAGTGCATGGTGCACATTTCCCGGTATCTATGATCCGAGCTTCACTCCGCTAAACAG 1738
QY 191 IuGluArgAlaAsnAsnValIleIysGlyThrIysGlnIleGlnIleValIleGlnIleIle 211
Db 1739 GCTCTCGGCCAACAGTGTCAATCAAGGCAACAAAGAAACAGGTGAGCAGATCATCA 1798
QY 211 YAspIle----- 213
Db 1799 AGGATATCAGGATATATGATATGATGCTTAAAGTCCCTTGGTCTAAGGTGACCCAGTG 1858
QY 213 ----- 213
Db 1859 CAACCTAAACAAATTAATTAATTAATTTGGTAAATATATACATATACAGAACAT 1918
QY 213 ----- 213
Db 1919 ATTTGTTAACCGGTGACTTAGAAGTCTGCATGATGTTGGACAAATTGACATTCGATA 1978
QY 213 ----- 213
Db 1979 TACAGTACCGCTCACTTGATGAGACTCCACAAAGAACTTAAACTTAAAGCTTAA 2038
QY 213 ----- 213
Db 2039 GCAACTATTCGTAGCTATGATGATTTGTGTGACATGATGTTGAAGATCTAGATTAACT 2098
QY 213 ----- 213
Db 2099 GGTGAAGAAATANGTTCACTAGCATTAATCATTAACAGAAAGCTATGCTTATGTA 2158
QY 214 ----- 230
Db 2159 GCTAAATGAACAGGAGTTTAAAGAGAGAAACAAAGTGAACAGATTAAGTGTGTGGA 2218
QY 230 hrrAlaAsnThrGluArgTYrSerAsnLeuValIleGlyLeuAsnAspThrMetGluAsn 250
Db 2219 CTGCAAAACATGAAGATATGATGATGTGCGCTGCTCAACGACACGATGAGATC 2278
QY 250 euleuAlaIleValAspArgAsnGlnIleGlnIleSerProSerThrIleuTYrAlaIle 270
Db 2279 TACTGCATCTGTGGAACAAGACGAGCGAGATATCACTCAACATATATGCTATTTG 2338
QY 270 laCyseValMetGluAsnValProPheIleAsnGlySerProGluAsnThrPheValPro- 289

```

Db      2339 CCTGTGTCAATGAGGGGGGCGCTTCATCATGGAGAGCCCGCAACACCTTTGTGCTG 2398
QY      289 -----
Db      2399 GTGCGGTGTGTGTGTGTGTGAAAAAGCTTCAAGGTGTTCATTTCTGTCCAAAGTTTC 2458
QY      290 -----
Db      2459 ATGGTGTGTATTTCCGTTCCAGGCTTATTTACCTGTGTGATGTTGCTAGGCGCTGATT 2518
QY      293 AAPPLeuAlaIleAlaArgAsnThrLeuIleGlyIleAspAspPheLysSerGlyGlnThr 312
Db      2519 GATCTTGTCTATTAATAAAACACACTGCTGATTGGTGTGAGCACTTCAAGAGTGAACAGACC 2578
QY      313 LysMetLysSerValLeuValAspPheLeuValGlyIleGlyIleLys----- 328
Db      2579 AAGATGAATCTGTCTGTGTGCAATTTCTTGTGTGGTGGGAAT-AAAGGTGGAAACCTA 2637
QY      328 -----
Db      2638 GTATCTCTCTCTATTAAGATGAAGTGTTTTGTGGCAATGACGTTATTGCATTAACCTC 2697
QY      329 -----
Db      2698 TTCTATATTTTCATTTTCATGCAAGCCCACTCAATCGTAGCTTCAACCACTTGGGAATC 2757
QY      341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnLysSerLysSer 360
Db      2758 AACGATGGCATGAACCTGCTGCTGCCCTTCAACATTCAGGTCCAAAGAGATCTCCAAAGAC 2817
QY      361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyGlnIlePro 380
Db      2818 AACGTGTGTGATGACATGCTCTCGAGCAATGCAATGCCATCTCTATAGACCCGCGCAAGATCC 2877
QY      381 AspHisValValValIleLys----- 387
Db      2878 GATCATGCTGCTTGTATCATCAGGCTGTGTAGCTGATCTTTCACCTGCTTAAAGTTGACAT 2937
QY      388 -----
Db      2938 ATGCAAGGACGATTTACATGTAACCTGTGCACTCTTTTGTGCAATAGTGTGCGTAACTGTG 2997
QY      393 GlyAspSerLysArgAlaMetAspGlyTyrThrSerGlnIlePheMetGlyValLysAsn 412
Db      2998 GGAGGACAGCAAGAGGCTATGACAGATGACATGATCTTCATGCGCGCAAGAAC 3057
QY      413 ThrIleValLeuHisAsnThrCysGlnAspSerLeuLeuAlaIleProIleIleLeuAsp 432
Db      3058 ACCATGTGTGTGACACACCTGTGAGAGCTGCTGCTGCGCACCTATCATCTCTTAT 3117
QY      433 LeuValLeuLeuAlaGlyLeuSerThrArgIleGlnPheLysAlaGlyLysAsnGlu----- 450
Db      3118 CTGGTGTCTTGTGTGAGCTGACACCGAGATCCAGATCGAAAGCTGAGAGGAGAGGTAAAGA 3177
QY      450 -----
Db      3178 GCCCCCCAAGTATTAACCTGAAGACAGCTGACGCTAGGTGATATAGCACTTTTAATA 3237
QY      451 -----
Db      3238 CCTTGTGTGTCTCTCTTATGACAGACAAATTCACCTCTTCACCGGTGCGCACCAATC 3297
QY      463 LeuSerTyrLeuThrLysAlaProLeu----- 471
Db      3298 CTGAGCTACCTCAACCAAGGCAACCTGTGAAGCTTTTCTCTGATCCCGGCATCACTG 3357
QY      472 -----
Db      3358 CACTGCGTTTGTCTCAATCCAGCACTGATGCTCTCTCGAAACTGAACAACAGGTT 3417
QY      473 ProProGlyThrProValAlaAsnAlaLeuSerLysGlnArgAlaMetLeuGlnLysIle 492
Db      3418 CCCCCTGGACACACCGGTGTGAACGCTCTGCGCAAGACAGACGCGATGCTGGAAGAACATC 3477

```

```

QY      493 MetArgAlaCysValGlyLeuAlaProGlnLysAsnMetIleLeuGlyTyrLys 510
Db      3478 ATGAGGGCGCTGCTGGGCTGCGCCAGAGAAACAATGATCTGAGTACAAG 3531

RESULT 5
US-09-677-064-15
/ Sequence 15, Application US/09677064
/ Patent No. 6291224
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
/ APPLICANT: Beach, Larry R.
/ TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
/ FILE REFERENCE: 0706D
/ CURRENT APPLICATION NUMBER: US/09/677,064
/ CURRENT FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: 60/055,446
/ PRIOR FILING DATE: 1997-08-11
/ PRIOR APPLICATION NUMBER: 60/055,526
/ PRIOR FILING DATE: 1997-08-08
/ PRIOR APPLICATION NUMBER: 60/053,944
/ PRIOR FILING DATE: 1997-07-28
/ PRIOR APPLICATION NUMBER: 09/118,442
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 15
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-677-064-15

Alignment Scores:
Pred. No.: 1.16e-218 Length: 3546
Score: 1823.00 Matches: 446
Percent Similarity: 40.86% Conservative: 28
Best Local Similarity: 38.45% Mismatches: 36
Query Match: 69.29% Indels: 652
DB: 3 Gaps: 9

US-10-718-952-12 (1-510) x US-09-677-064-15 (1-3546)
QY      1 MetPheIleGlnAsnPheLysValGlnSerProAsnValLysTyrThrGlnIle 20
Db      60 ATGTTTCATCGAGACTTCCGGTGTGAGAGCCCCACGTCGCGTACGGGCCGACGAGATC 119
QY      21 GlnSerValTyrAsnTyrGlnThrThrGlnLeuValHisGlnAsnArgAsnGlyThrTyr 40
Db      120 GAGTCGAGGTACCGGTACGACACGACGAGCTGTGTACACGAGGCGCAAGGACGGCGCTCA 179
QY      41 GlnThrIleValLysProLysSerValAsnTyrGlnPheLysThrAsnThrHisValPro 60
Db      180 CGCTGGGTCTGCGCCCAAGTCCGTAAAGTACATTCGAGACCAAGACCGCGTCCC 239
QY      61 LysLeuGly----- 63
Db      240 AAGCTCGGTATGTAGAGATGACAGGCGCCCTAGCTCACTCTGTGAACCTCTCTCTC 299
QY      63 -----
Db      300 CCGTGTCACTCAATCTCCGTGAGATGAACATGTGCGGTTCCCTCTAATCTTAA 359
QY      63 -----
Db      360 TGAATCTTACTGCTTTCCTGGAAGACGACCGTGTAAATTGTGACAGCTACGACAC 419
QY      63 -----
Db      420 ACTTGCCCATCCGAGTGCCTCAATTCAGCTCGATTGAATTGATTCGATTCGATTCGCTT 479
QY      63 -----

```


Db	1559	GGTGAACCCAGACGACACTTGTGTTCCGAGAGCGTGGGACATTAGCAAACTGAACCTCGGCGCA	1618
Qy	151	palamecAlaargAlaLysValPheasp11easpLeuGlnLysGlnLeuArgProTyrme	171
Db	1619	CTCCATGACCAAGGGCCCAAGGTGCTGGATATTGACCTTCAGAAAGCAAGCTCAGGCGCTTCAT	1678
Qy	171	tGluSerMetValProleu-ProGlyLylLeTyrAspProaspPhe11a1a1aasnGln	191
Db	1679	GGAGTTCATGATGTCACCTCCCGGTATGATATGATCCGAGCTTCATGCGGCTTAAACAGG	1738
Qy	191	luGluArgAlaasnAsnVal11elysGlyThrLysGlnGlnGlnValGlnGlnlelel	211
Db	1739	GCTCTCCGCGCAACAGTGTATCAAGGGCACCAAGAAAGAACAGGTGAGCACATCTCA	1798
Qy	211	ysaap11e-----	213
Db	1799	AGATATATCAGGTATATGATATGATGATCTAACGNGCCTTGCTGTAAGTGACCCAGTG	1858
Qy	213	-----	213
Db	1859	CAACCTAAAACAATAAATACTACTATGAATTGGTAATAATATACATATACGAACAT	1918
Qy	213	-----	213
Db	1919	ATTGTTTAACCGGTGCACCTTAGAAGTCTGCATGATGTTGGACAAATTGCATTCGATA	1978
Qy	213	-----	213
Db	1979	TACAGTACCGGTCTCACTTGATGATGAGCTCCACAAAGACTAAAACTAGAAAGCTTAA	2038
Qy	213	-----	213
Db	2039	GCAACTATTCGTAGCTAATGATGATATTGGTGGACATGTTGAAAGTCTAGATTAAAGT	2098
Qy	213	-----	213
Db	2099	GGTTGAAGAAATATGTTCTCACTACCATTAAGTATCCATTACAGAAAGCTATAGCTTATGTA	2158
Qy	214	-----LysAlaPheLysGlnLalThrLysValAspLysValValLeuTyrPT	230
Db	2159	GCTAATGAACAAGGAATTTTAAGAGAAAGAAACAAGTGACCAAGATGATTGTGTGTGGA	2218
Qy	230	hrAlaasnThrGluArgTyrSerAsnLeuValValGlyLeuAsnAspThrMetGluAsnL	250
Db	2219	CTGCAAAACACTGAAAGGTATAGCAATGTGTGCGCTGCTCCAACGACACATGAGATC	2278
Qy	250	euleuAlaValAspArgAsnGlnLalThrLysSerProSerThrLeuTyrAla11ea	270
Db	2279	TACGTGGATCTGTGGACAAAGACAGGCGGAGATACATCAACACTATATGCCATTG	2338
Qy	270	laCyValMetGluAsnValProPhe11easnGlySerProGlnAsnThrPheValPro-	289
Db	2339	CTGTGTGATGAGAGGGGTGCCCTTCATCAATGGAGAGCCCCAGAACACTTGTGTGCTG	2398
Qy	289	-----	289
Db	2399	GTGCGGTGTGGTGTGTTGCAAAAAGCTTCATGATGTTGCAATTCTGTTCCAAAGTTTC	2458
Qy	290	-----GlyLeu11e	292
Db	2459	ATGCGTGTGTAATTTCCGTTCCAAAGCTTATATATACCTGTTGCATGTTCCGATGGCGATTT	2518
Qy	293	AspLeuAla11leAlaArgAsnThrLeu11eGlyValAspAspPheLysSerGlyGlnThr	312
Db	2519	GATTTTCTATAAAAAACAACCTGCTTGAATGGTGGTGACACACTTCAAGAAGTGCACAGACC	2578
Qy	313	LysMetLysSerValLeuValAspPheLeuValGlyAlaGly11leLys-----	328
Db	2579	AAAGTGAATATCTGCTTGTGTCATTTCCCTGTTGGTGTGCTGAAT-AAAGGTGGGAACCTA	2637
Qy	328	-----	328
Db	2638	GTATCTCTCTTCTAATTAAAGATGAAGTGTTTTTTTGGCAAAATGACGTTATTTGCAATAACTC	2697

```
QY 359 -----ProthSerIleValSerTyrAsnHisIleuGlyAsn 340
Db 2698 TTCTATATTTTCATTTTCATGCGACCCACCTCAATCGTAGCTACCAACCTTGGAAGC 2757
QY 341 AsnAspGlyMetAsnLeuSerLysProGlnThrPheArgSerLeuGlnIleLeuTyrSer 360
Db 2758 AACGATGGCATGAACCTGCTGCTCCCTTCAACATTCAGGTCGACGAGATCTCCAGAGCC 2817
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGluHisPro 380
Db 2818 AACGTGGTGAATGAATGGTCTCGAGCAATGCCATCTCTAATGAGCCCGGCGAGATCCC 2877
QY 381 AspHisValValIleLys-----TyrValProTyrVal 392
Db 2878 GATCATGTCGTGTGATCATGAGGTCGTGTAGCTGATCTTACCTCGTTAAAGTTGACAT 2937
QY 388 -----TyrValProTyrVal 392
Db 2938 ATGCAAGCGAGATTATCATTTGAACTGTCACTCTTTGTTGACGATGTCGCGTACGTG 2997
QY 393 GlyAspSerLysArgAlaMetAspGluTyrThrSerGluIlePheMetGluValLysAsn 412
Db 2998 GGAGACAGCAGAGGCGCTATGAGACAGTACACCTCAGAGATCTTCATGGCGCGCAAGAC 3057
QY 413 ThrIleValIleuHisAsnThrCysGluAspSerLeuLeuAlaAlaProIleIleuAsp 432
Db 3058 ACCATCGTGTGCACACAACCTGTGAGGACTGCTCTCGCGCACTATCATCTCTGAT 3117
QY 433 LeuValIleuLeuAlaGluLeuSerThrArgIleGlnPheLysValGluLysGlu----- 450
Db 3118 CTGGGTCTTGTGCTGAGCTCAGACCAAGATCCAGCTGAAAGCTGAGGAGAGGTAAAG 3177
QY 450 ----- 450
Db 3178 GCCCCCCAAGTATTAACTGAAAGCAGCTCAGCGCTAGGTGATAGTACACTTTTAATA 3237
QY 451 -----GlyLysPheHisSerPheHisProValAlaThrIle 462
Db 3238 CCTTGTGGTGTCTCTTATATGACAGACAAATTCACCTCTTCCACCCGGTGGCACATC 3297
QY 463 LeuSerTyrLeuThrLysAlaProLeu----- 471
Db 3298 CTGAGCTACCTCACCAGGACACCCCTGTAAAGCTTTCTCTCGATCCCGGACATCTG 3357
QY 472 -----Val 472
Db 3358 CACTGCGTTTCTTCATCCAGCACTGATGCTCTCTCGAAACTGGAACAACAGTT 3417
QY 473 ProProGlyThrProValValAsnAlaLeuSerLysGlnArgAlaMetLeuGluLysIle 492
Db 3418 CCCCTGGACACACCGGTGTGAACGCTCTGCGCAAGCAGACGCGATGTGAGAAACATC 3477
QY 493 MetArgAlaCysValGlyLeuAlaProGluAsnMetIleLeuGluTyrLys 510
Db 3478 ATGAGGGCGCTGCTGGCTGGCCCAAGAGACAACATGATCTCTGAGTAAAG 3531

RESULT 6
US-09-118-442-14
; Sequence 14, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
```

```
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-14

Alignment Scores:
Pred. No.: 1,55e-218 Length: 3546
Score: 1822.00 Matches: 445
Percent Similarity: 40.86% Conservative: 29
Best Local Similarity: 38.36% Mismatches: 36
Query Match: 69.25% Indels: 652
DB: 3 Gaps: 9

US-10-718-952-12 (1-510) x US-09-118-442-14 (1-3546)
QY 1 MetPheIleGluAsnPheLysValGluSerProAsnValLysTyrThrGluThrGluIle 20
Db 60 ATGTTTCATCGAGAGCTTCCGGTCGAGAGCCCAAGCGGTACCGGCGGCGGAGATC 119
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 120 GAGTCGAGATACCGGTATGACACACAGAGAGCTGTACAGAGGCGCAAGAGAGCGGCTCA 179
QY 41 GlnTyrIleValLysProLysSerValAsnTyrGlnPheLysThrAsnThrHisValPro 60
Db 180 CGTGGTGTGCTCGCGCCCAAGTCCGTCAAGTAACTTCGAGACAGAAACCGCGTCCCC 239
QY 61 LysIleuGly----- 63
Db 240 AAGCTCGGTATGTAGCGATGACAGGCGCCPAGCCTCACTCTGTGAACCTCTCTCTC 299
QY 63 ----- 63
Db 300 CCGTGTCAATCAATCTCCGTGAGATCACTGTCGGGCTTCCTCTTAATCTTAA 359
QY 63 ----- 63
Db 360 TGAAATCTTACTGCTTTCCTGTAAGAGAACCGTCTGAATTTGTGACAGTACGACAC 419
QY 63 ----- 63
Db 420 ACTGCCCATCCGATGCGTCAAAATCAGCTGATTTGAATTCGATTCGATGTCCTT 479
QY 63 ----- 63
Db 480 TTCCATATTGATCATCTCCGCTACTGTGCAATGATTACAGAAACGTCCTTTCTCT 539
QY 63 ----- 63
Db 540 TGAATTTGTCTAGGCTTTTGTCTGTGACGAGAGCTGTATCAATTTGTTTATGTA 599
QY 63 ----- 63
Db 600 AGATCAATTCAGAGAGGAGAGATGAGAGAGACAGAACTATTACATAGCAAAATTGA 659
QY 63 ----- 63
Db 660 TACTAGATTATCGGACAGTGTGATACGAGGCAATCTGCATCTCGAACCCTTTGTT 719
QY 63 ----- 63
Db 720 TAATTCCTGTTCTATGATGATGTGCTTACGACAGCTGCTGTGTGTATGATGTCAGGC 779
QY 63 ----- 63
```

Db 780 TGTCAAGCCGCTTGTCTGTGTCGAGAGATGATGCCAACTTTTCTGTTCTGTGTGTCAGG 839
QY 64 ValMetLeuValGlyTTPGlyGlyValAsnAsnGlySerThrLeuThrGlyGlyValIleIle 83
Db 840 GTGAATGCTTGTGGGGGTGGGAGGCAACACGGGTCTCAGCTGACGGCTGGGTCAATGGCC 899
QY 84 AsnATGGLu----- 86
Db 900 AACAGGAGTGTAGTACTTAATTTGTCCTATATTGCTTCCGTTGTTTTCAGTTATTA 959
QY 87 -----AspIleSerTyrPalaThr 92
Db 960 ATGCCTAACAGAGAACTGAATTTTGTGTGTTGTTTCAAGGGATCTCATGGCCGAC 1019
QY 93 LysAspLysIleGlnGlnIleAsnTyrPheGlySerLeuThrGlnIleSerAlaIleArg 112
Db 1020 AAGGACAAAGTGTGACAGCCCACTACTACGGCTC-CTCACCCAGGCTCCACATCAGCA 1078
QY 113 ValGlySerPheGlnGlnGlyGlnGlnIleTyrPheProPheLysSerLeuLeuProMet 131
Db 1079 GTGGGAGCTACAAAGGGAGAGATCTATGGCGCTCAAGAGCCTCCCTCCCATGTGA 1138
QY 131 ----- 131
Db 1139 ATCTATTATAGACTTGACTAATACTCTCTTTTACTGAACCAACATACATACAAAG 1198
QY 131 ----- 131
Db 1199 CATATTTCCCTAAGGTCAGTGTGATGTATATAATGAACTGTCTTTCAGGCCAGTGTCT 1258
QY 131 ----- 131
Db 1259 TCAAGTAAAGGAAAGTTAATCATATTGGGTTGAAAAAACAAGGTCTAATTTTGAAAG 1318
QY 131 ----- 131
Db 1319 GAAAGTTAACTTAGCATATAGAAAAGGGAGAACACTGTAAAGAAAGTCTGAACAATC 1378
QY 131 ----- 131
Db 1379 GACTGGTCTGCCATGTGTGATCTCTACTTGCAAGTCAAAAGGTTCTGTGTAGCCAA 1438
QY 131 ----- 131
Db 1439 AGGTTCCAGCATCTTTGATTAACACTGTGAGATTTAGACATGTGTCTAACTGTTGCA 1498
QY 131 ----- 131
Db 1499 GATTGGCAGACTCGGTTTGTATCTTTCATGACCAAGTGTAAACGTGTTTCA 1558
QY 132 -ValAsnProAspAspIleValPheGlyGlyTTPAspIleSerAsnMetAsnLeuAlaAs 151
Db 1559 GGTTGAACCCAGACGACATGTGTGCGAGGCTGGGACATTAGCAACATGAACTGGCCGA 1618
QY 151 PalMetLeuAlaArgAlaLysValPheAspIleAspLeuGlnLysGlnLeuLeuArgProTyrMe 171
Db 1619 CTCCTAGACCAAGGCTCTGTGATTTGACCTGCAAGACAGCTCAAGCCCTCAACT 1678
QY 171 GlnSerMetValProLeu-ProGlyIleTyrAspProAspPheIleAlaAsnGlnG 191
Db 1679 GGAAGTCACTGTGCCACTTCCCGGATCTATGATCGGACTTTCATCGCGGCTAAACAGG 1738
QY 191 LysGlnArgAlaAsnAsnValIleLysGlyThrLysGlnGlnGlnValGlnGlnIleIleL 211
Db 1739 GCTCTGCGCCCAACAGTGCATCAAGGGCAACCAAGAAAGAACAGGTGGAGCATCATCA 1798
QY 211 YAspIle----- 213
Db 1799 AGGATATACGGTATATGATATGATGATGTAACGTGCTTGTGTAAAGTGCAACCCAGTG 1858
QY 213 ----- 213
Db 1859 CAACCTAAAAAATAATTAATACTATGAAATTTGGTAATATACATATACATACAGACAT 1918

QY 213 ----- 213
Db 1919 AATGTTTAAACGGTGCACTTAGAGACTGTGCATGTATGTGTGACAATTGGACTTCGATA 1978
QY 213 ----- 213
Db 1979 TACAGTACCGCTCACTTGATGAGAGACTCCACAAAGAACTAAAACACTAGAAAGCTTAA 2038
QY 213 ----- 213
Db 2039 GCAACTATTGTAGCTAATGATGATATTGTGTGACATGTGTGAAGATCTAGATTAACT 2098
QY 213 ----- 213
Db 2099 GGTGAAGAAATATGTTCTCACTAGTATATGTAATCATTCATTACGAAGCAATGCTTATGTA 2158
QY 214 -----LysAlaPheLysGlnAlaThrLysValAspLysValValIleLeuTyrPT 230
Db 2159 GCTAATGAACACGGAGCTTTAGGGAGAGAACAAAGGTGACAAAGATGTTGTGTGTGA 2218
QY 220 hrAlaAsnThrGlnArgTyrSerAsnLeuValGlyLeuAsnAspThrMetGluAsnL 250
Db 2219 CTGCAAAACACTGAAGATATAGCAATGTGTGCGTGTCTCAACGACACGATGAGAAATC 2278
QY 250 euLeuAlaAlaValAspArgAsnGlnAlaGlnIleSerProSerThrLeuTyrAlaIleA 270
Db 2279 TACTGCACTGTGTGACAAAGACGAGGAGGTATCACCATCAACACTATATGCTATTG 2338
QY 270 laCysValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValPro- 289
Db 2339 CCTGTGTCATGGAGGGGGGCGCTTCATCAATGGAAGCCCCAGAACACCTTTGTGCTTG 2398
QY 289 ----- 289
Db 2399 GTGCGGTGTGTGTGTGTGCAAAAGCCTCATGTGTGATTTCTGTCCAAAGTTTC 2458
QY 290 -----GlyLeuIle 292
Db 2459 ATGATGTGTATTTCTGTCCAAAGCTTATTAACCTGTGACATGTCTGATGGGCTGATT 2518
QY 299 AspLeuAlaIleAlaArgAsnThrLeuIleGlyGlyAspAspPheLysSerGlyGlnThr 312
Db 2519 GATCTTGCTATAAAAACACTGCTTGATGTGTGATGACATCAAGATGTGACAGACC 2578
QY 313 LysMetLysSerValLeuValAspPheLeuValGlyAlaGlyIleLys----- 328
Db 2579 AAGATGAATCTGTGTGTGATTTCTGTGTGTGCTGGAAAT-AAAGGTGGAACCTTA 2637
QY 328 ----- 328
Db 2638 GTATCTCTTCTATTAAGATGAAGTGTGTTTGGCAAAATGACGTATATGCAATAACTTC 2697
QY 329 -----ProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db 2698 TTCTAATATTTTCATTTTCATGACGACCCACCTTAATGTGAGGTAAACCACTTGGGAAC 2757
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlnIleSerLysSer 360
Db 2758 AACGATGGCATGAACCTGTCTGCCCTTCAACATTCAGGTCCAAAGAGATCTCCAAAGCC 2817
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGlnProGlyGlnHisPro 380
Db 2818 AACGTGTGTGATGACATGTGTCTCGAGCAATGCAATCTCTATGAAGCCCGGCGAGCATCCC 2877
QY 381 AspHisValValValIleLys----- 387
Db 2878 GATCATGTGTGTGATCATCAAGGTCTGTAGCTGATCTTCACTGCTTAAAAAGTTGACAT 2937
QY 388 -----TyrValProTyrVal 392
Db 2938 ATGCAAGCAGATTTTACATGAAACTTGACACTCTTTTGTGCAATATGTGCGGATACGG 2997

QY 393 G1yAspserlysaArgAlaMetAspGluTrpThrSerGluLeuPheMetGlyGlyLysAsn 412
Db 2998 GGAGACAGCAAGAGGCTATGACGAGTACACCTCAGAGATCTTCATGGCGGCAAGAAC 3057
QY 413 ThrileValleuHiasnThrCyseGluAspSerLeuValAlaProIleleuAsp 432
Db 3058 ACCATCGTGCTGCAACAACCTGTAGAGCTGCTCCGCGCACCTATCATCTTGAT 3117
QY 433 LeuValleuValAlaGluLeuSerThrArgIleGlnPheValAlaGluAsn----- 449
Db 3118 CTGGTCTCTTGCTGAGCTGACGACAGAGATCCAGTGAAGCTGAGGAGGGGTAA 3177
QY 449 ----- 449
Db 3178 GCGCCCAAGTATTAACCTGAAGACGCTGACGCTAGTATAGACTTTTAA 3237
QY 450 -----GluGlyLysPheHisSerPheHisProValAlaThrIle 462
Db 3238 CCTTCGTGTCTCTTATGACGACAAATTCACCTCTCCACCCGGTGGCCACATC 3297
QY 463 LeuSerTrpLeuThrLysAlaProleu----- 471
Db 3298 CTGAGCTACTCACCAGGACCCCTGTAGCCTTTCTCTGCAATCCGCGCATCACTG 3357
QY 472 -----Val 472
Db 3358 CACTGCGTTTGCTTCATCCAGCACTGATGCTCTTGAACCTGAACAAGGTT 3417
QY 473 ProProGlyThrProValValAsnAlaLeuSerLysGlnArgAlaMetLeuGluAsnIle 492
Db 3418 CCCCCGGACACCCGGTGTGAACCTCTGCGCAAGACGCGCATGCTGAGAGAACATC 3477
QY 493 MetArgAlaCyseValGlyLeuValProGluAsnAsnMetIleLeuGluThrLys 510
Db 3478 ATGAGGCGCTGCTGGCTGGCGCCCAAGAACATGATCTGAGTACAG 3531

RESULT 7

US-09-677-064-14
Sequence 14, Application US/09677064
Patent No. 6291224
GENERAL INFORMATION:
APPLICANT: Martino-Calt, Susan J.
APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.
TITLE OF INVENTION: Genes Controlling Phosphate Metabolism in
FILE REFERENCE: 0706D
CURRENT APPLICATION NUMBER: US/09/677,064
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/055,446
PRIOR FILING DATE: 1997-08-11
PRIOR APPLICATION NUMBER: 60/055,526
PRIOR FILING DATE: 1997-08-08
PRIOR APPLICATION NUMBER: 60/053,944
PRIOR FILING DATE: 1997-07-28
PRIOR APPLICATION NUMBER: 09/118,442
PRIOR FILING DATE: 1998-07-17
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 3546
TYPE: DNA
ORGANISM: Zea mays
US-09-677-064-14

Alignment Scores:

Pred. No.: 1.55e-218 Length: 3546
Score: 1822.00 Matches: 445
Percent Similarity: 40.86% Conservative: 29
Best Local Similarity: 38.36% Mismatches: 36
Query Match: 69.25% Indels: 652
DB: 3 Gaps: 9

US-10-718-952-12 (1-510) x US-09-677-064-14 (1-3546)
QY 1 MetPheIleGluAsnPheLysValGluSerProAsnValLysTrpThrGluThrGluIle 20
Db 60 ATGTTTCATCGAGAGCTTCGCGCTGAGAGCCGCCACGTCGGGTACGGCCGACGAGATC 119
QY 21 GlnSerValTrpAsnTrpGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTrp 40
Db 120 GAGTCGAGATACCGGTACAGACAGACGAGCTGTGACACGAGGCGCAAGAGCGGCTTCA 179
QY 41 GlnTrpIleValLysProLysSerValAsnTrpGlnPheLysTrpAsnThrHisValPro 60
Db 180 CGTGGGTGCTGCCGCCCAAGTCGTCAAGTACAACTTCGAGCCAGAACCGCGTCC 239
QY 61 LysLeuGly----- 63
Db 240 AAGCTCGGTATGTAACGATGACGAGGCGCTTACCTCACTCTGTGTAACCTCTCTC 299
QY 63 ----- 63
Db 300 CCGTGTCAATCAATCCTCCGTGAGATCAACTGTGCGGCTCCCTCTAAATCTTAA 359
QY 63 ----- 63
Db 360 TGAATCTTACTGCTTTCCTGGAAGAACCGTGTATTTGTGACAGTACGACAC 419
QY 63 ----- 63
Db 420 ACTTGCCATCCGATGCTGCAATCACTGATTTGAAATTCGATTGATGTCCTT 479
QY 63 ----- 63
Db 480 TTCCATATTGATCATCCTCGCTACTGTGCAATGATTAAGAAACGTCCTTTCTC 539
QY 63 ----- 63
Db 540 TGAATTTGTCTTAGCTTTTGTCTGTGACGAGTGTGATCAATTTGTTCANGTA 599
QY 63 ----- 63
Db 600 AGATCAATTCACAGCAGGAGATGACGACAGACAGCAACATCTACCTAGCAATTA 659
QY 63 ----- 63
Db 660 TACTAGATTAOTGGCAAGTGATGATACGGGCAATCTGCATCTGACCCCTTGT 719
QY 63 ----- 63
Db 720 TAATTCCTTCTTATGATGATGTGCTTACGTCAGCTGTGTGTATGATGTCAGGC 779
QY 63 ----- 63
Db 780 TGTCAACCGCTTGTCTGTCTGTCGACGATGATGCCAATTTTGTGTCGTGTCAGG 839
QY 64 ValMetLeuValGlyTrpGlyGlyAsnAsnLysSerThrLeuThrGlyValIleAla 83
Db 840 GTGATGCTTGTTGGGTGGGAGGACAGACAGGCTTCAAGCGGCTGAGGCTTATTC 899
QY 84 AsnArgGlu----- 86
Db 900 AACAGGAGTGAATGATTAATTTGTCTTATATTCCTTCCGTGTTTCAGTTATTA 959
QY 87 -----AspIleSerTrpAlaThr 92
Db 960 ATGCTTACAGAGAACTGAATTTTGTGTTGTTTTCAGGGGATTCATGCGCAGC 1019
QY 93 LysAspLysIleGlnGlnAlaAsnTrpPheGlySerLeuThrGlnAlaSerAlaIleArg 112
Db 1020 AAGGACAAAGTGCACAGCAACCACTACACGCTC-CTCACCCAGGCTCCACATCA 1078
QY 113 ValGlySerPheGlnGlyGluGluIleTrpAlaProPheLysSerLeuLeuProMet--- 131
Db 1079 GTCCGCACTCAACGCGGAGAGATCTATGCGCGTTCAAGAGCCTCTCCATGTA 1138

QY 131 ----- 131
Db 1139 ATCTATTATAGACTTGACTTAATCTCTCTTTTACTGAAACCAATACATATACAAAG 1198
QY 131 ----- 131
Db 1199 CATATTCCGTAAAGTGTAGTTGATGTATATAAATGAACCTGTCTTTCAGGCCAGTGTCTC 1258
QY 131 ----- 131
Db 1259 TCAAGTAAACGGAATGTATATCATTTGGTTGAAAAACAAGGTCTAAATTTTGTGAAG 1318
QY 131 ----- 131
Db 1319 GAAAGTTAACTTAGCATATGAAAGGGAGAGACTGTAAAGAAAGTGTGAAACATC 1378
QY 131 ----- 131
Db 1379 GACTCGGTCTGCATGTGTGATCTTCTGCAAGTCAAAAAGTCTGTGTAGCCCA 1438
QY 131 ----- 131
Db 1439 AGGTTCCAGCATCTTGTGATTACATCTGTGAGATTGACATGTGTCTAATCTGTTGCA 1498
QY 131 ----- 131
Db 1499 GATTCGCAGACTCGGTGTTTGTATTCTCTTTTCATGACCAAGTGTAAACGTGTTTCA 1558
QY 132 -ValAsnProAspAspIleValPheGlyValTyrPheAsnMetAsnLeuAlaAsn 151
Db 1555 GGTTGAACCCAGACGACATTTGTGTGGAGGCTGGGACATTTAGCAATATACTGGCCCA 1618
QY 151 PAlaMetAlaArgAlaValPheAspIleAsnLeuGlnValGlnLeuArgProTyrMet 171
Db 1619 CTCCTATGACCAAGGCGCAAGTGTGTGATATTGACCTGCAAGAACAGCTCAGGCCCTACAT 1678
QY 171 LeuIleuMetValProLeu-ProGlyIleTyrAspProAspPheIleAlaAlaAsnGlnG 191
Db 1679 GGAGTTCAGATGGTCCCATCTCCCGGTATGTATGATCCGACCTTCAGCGGCTAACCAAG 1738
QY 191 LuGIuArgAlaAsnValIleLeuGlyThrLeuGlnGluValGlnGlnIleIle 211
Db 1739 GCTCTCGCCCAACAGTGTATCATACAGGCAACCAAGAAAGACGGTGGAGCAGATCATCA 1798
QY 211 YAspIle----- 213
Db 1799 AGGATATACGATATATGATATGATGATGTAACGTGCTGTGTAAAGTGCACCCAGTC 1858
QY 213 ----- 213
Db 1859 CAACCTAAAACAATATAATACTACTATGAATTTGGTAATATATACATATCAGAGCAT 1918
QY 213 ----- 213
Db 1919 ATTGTTTAACCGGTGACTTAGAGTCTGATGATGTGTGACAATTTTGACATTCGATA 1978
QY 213 ----- 213
Db 1979 TACAGTACCGCTCACTTGTCATGAGAGATCTCAAAAGAACTAAACTACTGAAAGCTTAA 2038
QY 213 ----- 213
Db 2039 GCAACTATTCTGATGTAATGATGTATTGGTGACATGGTTGAAGATCTAGATTAACT 2098
QY 213 ----- 213
Db 2099 GGTGAAGAAATATGTTCACTAGTATATAGTATATCATTTACGAAGCAATGGCTTATGTA 2158
QY 214 -----LysAlaPheLeuGlyValAlaThrLeuValAspLeuValValLeuTyrP 230
Db 2159 GCTAATGAAACAGGAGTTTAGGAGAGAAAGAACAAAGTGGACAAAGATAGTGTGTGTGA 2218

QY 230 hrAlaAsnThrGluArgTyrSerAsnLeuValGlyLeuAsnAspThrMetGluAsnL 250
Db 2219 CTGCAAAACACTAAGATATAGTATGTGCGCTGTGCTTCAACGACAGATGAGAAATC 2278
QY 250 euLeuAlaIleValAspArgAsnGluAlaGluIleSerProSerThrLeuTyrAlaIleA 270
Db 2279 TACTGCATCTGTGACAAAGAACGAGCGGAGATACCATCAACACTATATGCAATTG 2338
QY 270 IeCyValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValPro 289
Db 2339 CCTGTGTATGAGAGGCGGTGCTTCATCATATGAGAGCCCCAGAACACTTTGTGCTTG 2398
QY 289 ----- 289
Db 2399 GTGCGGTGTGTGTGTGTTGCAAAAACCTCATGTGTGATTTCTGTCCAAAGTTTC 2458
QY 290 -----GlyLeuIle 292
Db 2459 ATGTGTGTATTTCTGTCCAAAGCTTATATACCTGTGTGATGTGCTAGGCTGATT 2518
QY 293 AspLeuAlaIleAlaArgAsnThrLeuIleGlyIleAspAspPheLeuSerGlyGlnThr 312
Db 2519 GATCTTGCTATATAAAACAACCTGCTTGTGATGTGTGAGCAGCTTCAAGAGTGGACAGACC 2578
QY 313 LysMetLysSerValLeuValAspPheLeuValGlyAlaGlyIleLeu----- 328
Db 2579 AAGATGAATCTGTCTGTGTGATTTCTTGTGTGTGCTGAAT-AAAGTGGAAACCTA 2637
QY 328 ----- 328
Db 2638 GTATCTCTCTTATTAAGATGAGTGTGTTTGGCAAAATGACGTATATGCAATATCTC 2697
QY 329 -----ProThrSerIleValSerTyrAsnHisLeuGlyAsn 340
Db 2698 TTCTATATTCTTATTTTCATGACGAGCCCACTCAATCGTAGCTAACACACTTGGGAAAC 2757
QY 341 AsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLeuGluIleSerLysSer 360
Db 2758 AACGATGGCATGAAACCTGTCTGCTCTTCAMAACTTCAGGTCCAGAGAGTCTCCAGAAC 2817
QY 361 AsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluProGlyValHisPro 380
Db 2818 AACGTGTGTGATGATGATGTCTCGAGCAATGCCATCTTATGAGCCCGGCGAGCATCCC 2877
QY 381 AspHisValValIleLeu----- 387
Db 2878 GATCATATGCTGTATCATCAAGGTCTGTAGTGTATCTTTCACCTCGTTAAAGTTGACAT 2937
QY 388 -----TyrValProTyrVal 392
Db 2938 ATGCAAGCAGATTATACATTGAAACTTGTCACTTTTGTGTGACATATGTGCGTACATG 2997
QY 393 GlyAspSerLysArgAlaMetAspGluTyrThrSerGluIlePheMetGlyValLysAsn 412
Db 2998 GGAGACAGCAAGAGGCTATGTAGACAGATACACTCAGAGATTTTATGTGGCGGCAAGAAC 3057
QY 413 ThrIleValLeuHisAsnThrCysGluAspSerLeuLeuAlaIleProIleIleLeuAsp 432
Db 3058 ACCATCGTGTGCAACAACCTGTGAGACTCGTCTCGCGGCACTATCATCTTGTAT 3117
QY 433 LeuValIleuLeuAlaGluLeuSerThrArgIleGlnPheLeuAlaGluAsn----- 449
Db 3118 CTGTGTCTTGTGCTGAGCTCAGCACCAAGATCCAGTGAAGAGCTGAGAGGAGGTAAAG 3177
QY 449 ----- 449
Db 3178 GCCCCCAAGTATTAACCTGAAGACAGCTGCAGCTAGGTGATATAGCACTTTAATA 3237
QY 450 -----GluGlyLysPheHisSerPheHisProValAlaThrIle 462
Db 3238 CCTTGTGTGTCTCTTATATGACAGAACAAATTCACCTCTTCCACCGGTGGCCACATTC 3297
QY 463 LeuSerTyrLeuThrValProLeu----- 471

QY 485 GlnArgAlaMetLeuGluIleuValMetArgAlaCysValGlyLeuAlaProGluAsn 504
DB 1516 CAAGAAGCCGCTTGAAGAAATTTTAAAGATTGTTGATTCGCTTCTCAAAACGAA 1575

QY 505 MetIleLeuGlu 508
DB 1576 CTAAAGATTCGAA 1587

RESULT 9
US-09-734-237B-74
; Sequence 74, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic gene derived from Saccharomyces cerevisiae myo-inositol
; OTHER INFORMATION: -1-phosphoric synthase, having numerous codons replaced with other
; OTHER INFORMATION: s encoding the same amino acids to reduce free energy of folding,
US-09-734-237B-74

Alignment Scores:
Pred. No.: 1,42e-164 Length: 1605
Score: 1389.50 Matches: 278
Percent Similarity: 69.66% Conservative: 87
Best Local Similarity: 53.05% Mismatches: 138
Query Match: 52.81% Indels: 21
DB: 4 Gaps: 8

US-10-718-952-12 (1-510) x US-09-734-237B-74 (1-1605)

QY 3 IleguaAnPheLyValGluSerProAsnValLySerThrGluThrGluIleGlnSer 22
DB 28 ATCATCTTCTGTTAAAGTTGTAAGTGAACAATGTAACAAAGACAAACGACGCGTCACT 87

QY 23 ValIlyrAsnThrGluThrGluLeuValHisGluAsnArgAsnGlyThrTyrglnTrp 42
DB 88 AATATCTCTTACGAAACCGCT--GTAAGTAACTAAACGCTTCGTCGTTTCGAT-- 141

QY 43 IleValLyProLySerValAsnThrGluIlePheLyThrAsnThrHisValPro--Lys 61
DB 142 ---GTTATCCGACGCTTGAAGCTACGATTCACAACTGGAATCTGAAGAAACGGAAAG 198

QY 62 LeuGlyValMetLeuValGlyTyrGlyGlyAsnAsnGlySerThrLeuThrGlyGlyVal 81
DB 199 CTGGGTATCATGCTGATCGCGCTGGGTGTAAACAACGCTCACTCTGTTGCATCTGTT 258

QY 82 IleAlaAsnArgGluAspIleSerThrPalaThrLysAspLysIleGlnGlnAlaAsnThr 101
DB 259 CTGGCAACAAACACAACTAGAAATTCACAAAGAAAGAGGTAAACAGCCGAACTAC 318

QY 102 PheGlySerLeuThrGlnAlaSerAlaIleArgValGly---SerPheGlnGlyGluGlu 120
DB 319 TTGAGTTCTATACCTACAGTCTTCTACTCGAAGCTGGGCAATGATGCTGAAGGTAAACGAC 378

QY 121 IleTyraIleProPheLysSerLeuLeuProMetValAsnProAspAspIleValPheGly 140
DB 379 GTTATCGCTCCGTTCAACTCTCTGCTGCGGATGATGATCTCCGAACGACTTGTGTTTCT 438

QY 141 GlyTPAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPheAsp 160
DB 439 GGTGGATATCAACAACCGCATCTGTACGAAGCAATGACGCTTCTCAAGTTCTCGAA 498

QY 161 IleAspLeuGlnLysGlnLeuArgProTyMetGluSerMetValProLeuProGlyIle 180
DB 499 TATGATCTGCAACAGCGCTGTAAGGCTAAGATGTCCTCTGTTAAGCACTGCGCTCATC 558

QY 181 TyraAspProAspPheIleAlaAlaAsnGlnGluGluValAlaAsnValIle----- 198
DB 559 TACTACCCGATTTATCGACCTAACCAAGCAAGCAAGTGTAAACAGTATCAACCTG 618

QY 199 -----LysGlyThr-----LysGlnGluGlnValGlnIleLys 211
DB 619 GACGAAAGGGTAAACGTTATCAACCGGTGTGTGTGAGTCACTCACTGAGGTATCCGTCGT 678

QY 212 AspIleLysAlaPheLysGlnAlaThrLysValAspLysValValLeuThrPThrAla 231
DB 679 GATATCCAGAACTTCAAGAGGAAACGACCTGCAACAAAGTTATGTTACTGTGACTGCT 738

QY 232 AsnThrGluArgTyrSerAsnLeuValValGlyLeuAsnAspThrMetGluAsnLeu 251
DB 739 AACACTGAACGTTTACGTAGAGTATCCCGGCTGTAAACGATATATGGAACCTGCTG 798

QY 252 AlaAlaValAspArgAsnGluAlaGluIleSerProSerThrLeuThrIleAlaCys 271
DB 799 CAATCTATCAAGACGACGACGAGAAATCGCTCCGTCACATCTCTCTGATCT 858

QY 272 ValMetGluAsnValProPheIleAsnGlySerProGlnAsnThrPheValProGlyLeu 291
DB 859 ATCCGTAAGAGGCTACCGTATCAACGCTCTCCGCAACACATTGCTACCGGCTG 918

QY 292 IleAspLeuAlaIleAlaArgAsnThrLeuIleGlyGlyAspAspPheLysSerGlyGln 311
DB 919 GTACAGCTGCTGAAACAGAAAGTACTTCACTGCTGTGACGATCTGAATCTGAGCCAG 978

QY 312 ThrLysMetLysSerValLeuValAspPheLeuValGlyAlaGlyIleLysProThrSer 331
DB 979 ACTAAATCGAAATCTGTACTGCGACAGATTCCTGTTGACGCTGTATCAACCGGTTT 1038

QY 332 IleValSerTyrAsnHisLeuGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThr 351
DB 1039 ATCGCTTCTTATACCACTCGGTGTAACAACGCTTCAACCTTCTCTCCGAACG 1098

QY 352 PheArgSerLysGluIleSerLysSerAsnValValAspAspMetValAsnSerAsnAla 371
DB 1099 TTCGCTTCTAAAGAAATCTCTAAATCTCTGTATCTGATCGACATCATCGCTTCTAAACGAC 1158

QY 372 IleLeuTyrgluPro-----GlyGluHisProAspHisValValIleLysTyrglyVal 389
DB 1159 ATCCGTGTACACGACAAACGCGGTAAAGATGATCACTGTATGTATCAAAATACATG 1218

QY 390 ProTyraValGlyAspSerLysArgAlaMetAspGluTyThrSerGluIlePheMetGly 409
DB 1219 AAACCGGTTGGATTTCTAAAGTCTTATGACGAATCTACTCTGAACCTAGTCTGGGC 1278

QY 410 GlyLysAsnThrIleValLeuHisAsnThrCysGluAspSerLeuLeuAlaAlaProIle 429
DB 1279 GGTACAAACCGATCTCTATACCAACAGTTTGTGAAGCTCTGCTGCTGCTACCCGCTG 1338

QY 430 IleLeuAspLeuValLeuLeuAlaGluLeuSerThrArgIleGlnPheLys----- 446
DB 1339 ATCATCGACCTGCTGTATATGATGTAATTCGTATCCCGTGTATCTTAACAAAGTTCAC 1398

QY 447 -----AlaGluAsnGluGlyLysPheHisSerPheHisProValAlaThrIleLeuSer 464
DB 1399 CCGGTTAAAGAAAGATGCTGGCAATTCGAAACATTCTACCCGTTCTACCTTCTGCTCT 1458

QY 465 TyrlleuThrLysAlaProLeuValProProGlyIleThrProValValAsnAlaLeuSerLys 484
DB 1459 TACTGGCTGAAGAGCTCCGCTGATCTGTCAGAGCTTCCACCGGTTAAACGCTGTAACAA 1518

RESULT 10

US-09-248-796A-3131
; Sequence 3131, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NOCLETIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3131
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-3131

Alignment Scores:

Pred. No.:	5.9e-164	Length:	1578
Score:	1384.50	Matches:	266
Percent Similarity:	71.29%	Conservative:	99
Best Local Similarity:	51.95%	Mismatches:	132
Query Match:	52.62%	Indels:	15
DB:	4	Gaps:	5

US-10-718-9552-12 (1-510) x US-09-248-796A-3131 (1-1578)

OY 9 GluSerProAsnValIysThrGluThrGluIleGlnSerValIyrAsnTyrgLunrh 28
Db 37 AAATGTCCAAATTCTGTCACTAAGACGCATCATTTATATACCAATTACTTATGAACAAC 96
OY 29 ThrGluLeuValHisGluAsnArgSasnGlyThrTYrGlnTrpIleValIysProlysSer 48
Db 97 TCT--GTTGTTGAAAAAAGATGCTAATGCT-----AAATTCATGTTACACCAACATGCT 147
OY 49 ValAsnTYrGlnPheIysThrAsnThriSVAlProIySLeuGLyValMetLeuValIely 68
Db 148 AGTGCACTTGAATTCAAAAGTTGATTAAAGTCCCTTAAGTGGTGTGTTATGTATGTCGCT 207
OY 69 TrpGIyGLYAenASngLYSerThrlEuThrgIyGLyVALIlEAlAsnArgLUaspIIle 88
Db 208 ATTGGCGGGAACACAGGACTACATTCATTTAGGTGCCACATTAGCACACAAACACAAATTT 267
OY 89 SerTPalArthrLyASPILysIleGInGlnIlaSenTYrPheGlySerLeuThrgIlnAla 108
Db 268 TCTTTTGAACAACAAGAAGCGTGTGTTAAACCATAATTCATGCTTCTGTTACTCAAGCA 327
OY 109 SerAlaIleArgValGly-----SerPheGInGlyGluGluIleTYrAlAProPhelYS 126
Db 328 TCCACTGTCCAANAATCGGTGTGATTAAGAAACCTGGTGAAGATGTTACGTTCCATTTAAC 387
OY 127 SerLeuLeuPROMetValAsnProAspAspIleValIPheGlyGlyTrpAspIleSerAsn 146
Db 388 TTCATTGTTCCAAATGTTAAACCCAAATGATTTGGTGTGATGAGTGGGATTAATTAAGTGT 447
OY 147 MetAsnLeuAlaASpAlAMerAlaArgAlaIySValIPheAspIleASPLeuGlnIynSgIn 166
Db 448 TTGCATTTAGAACCAAGCAATGAAGAGAGCTAAAGCTTGAGATGTTACTTTACAAAAGAG 507
OY 167 LeuArgProTYrMetGlusErMerValProLeuProGlyIleTYrAsPProAspPheIle 186

Ds	508	TTGATATCATTCCTTGGAGAACAGAACCTTTGGATCAATCTATCACTGATTTTCATTT	567
Qy	187	AlAlaIaasnnglmgluIaArgAlaAsnAsnValIleLys-----	199
Ds	568	GCTTTGAACCAAGATGAAACGTGGCCAAACAATGTTTTTAAACAAGTCAATGCTGAAGTTAA	627
Qy	200	---GlyThrLysglnglmglnValAlnglnmlleIleLysAspIleLysAlaPheLysglu	218
Ds	628	ACTGATATATAATGAGCGCGACGTGAAAAAATACGAAAMAAATATACAGAGTTTCAAGGCC	687
Qy	219	AlaThrLysValAspLysValIleValIleuTriPheAlaAsnThrGluArgLysSerAsn	238
Ds	688	AAAAACGAATTAGATTAAAGTTATTATTATTATTTGACAGCTCTAATACTGAAACATATGCTGAT	747
Qy	239	LeuValValGlyLeuAsnAspThrMetGluAsnLeuLeuAlaIleValAspArgAsnLys	258
Ds	748	GTTTGGCAAAATGTCATATGACACGCTGACACAACTGATTCAAAGATATTAAAGAATCAC	807
Qy	259	AlaGluIleSerProSerThrLeuThrValAlleAlaLysValMetGluAsnValProPhe	278
Ds	808	GAACAATAATGGCTCCAAAGTACTGCTCTTGGCCGTGCTTCTATCGTGAAGAAAGTTCCATAT	867
Qy	279	IleAsnGlySerProGluAsnThrPheValProGlyLeuIleAspLeuAlaIleAlaArg	298
Ds	868	ATTATAGGTTACACAAACAAACACATTTGTTCCGAGTCTTATTGAGTTAGCTGAAAAATAC	927
Qy	299	AsnThrIleuIleGlyGlyAspAspPheLysSerGlyGlnThrLysMetLysSerValLeu	318
Ds	928	GATCATTCATATGGTGGTGATGATTTTCAAGTCAGGTCAACAAATAATTAATCAGTGTTA	987
Qy	319	ValAspPheLeuValGlyAlaGlyIleLysProThrSerIleValSerLysTranHisLeu	338
Ds	988	GCTCAATTCCTGGTCGATGCTGATTAACCAACTTCTTATTCCTTAAATATCATCTTG	1047
Qy	339	GlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGluIleSer	358
Ds	1048	GGTAAACAATGACGGTTTCAATTATTCATCCCAAAACAATTATGATTTAAGAAATTTCC	1107
Qy	359	LysSerAsnValValAspAspMetValAsnSerAsnAlaIleLeuTyr-----GluPro	376
Ds	1108	AAACAATCTGTTGTTGATGATATTAATGAATCCACAACAATTAATGTATCAACAAGAATCT	1167
Qy	377	GlyGluHisProAspHisValValIleLysTyrValProTyrValGlyAspSerLys	396
Ds	1168	GGTACCAAGTTGATCACTGATGCTATTAATAATCACTGCAGCTGTTGGTGATTTCTAAA	1227
Qy	397	ArgAlaMetAspGluTyrTrpHisSerGluIlePheMetGlyGlyLysAsnThrIleValLeu	416
Ds	1228	GTTGCCAATGATGAAATCTATTCAGAAATTAATGTTGGTGTGTCTCAACAATAAATGATATT	1287
Qy	417	HisAsnThrCysGluAspSerLeuLeuAlaAlaProIleIleLeuAspLeuValLeuLeu	436
Ds	1288	CACAAATGTTTGTGAAGATTCATCTGCTTGCTACACCAATGATTTATGACTTGGTGGTGGCT	1347
Qy	437	AlaGluLeuSerThrArgIleGlnPheLysAlaGluAsnGluGlyLysPheHisSerPhe	456
Ds	1348	ACTGAATAATTTGCAACTAAGTTCAAGTCAAAGAGTCTCGTGAATACTGATTTGATGATTTG	1407
Qy	457	HisProValAlaIleThrIleLeuSerTyrLeuThrLysAlaProLeuValProProGlyThr	476
Ds	1408	TATCTCTGTGTAGATTATTTGTTACTTACTGCTCAAGGCTCCATTTAGCAACACAGGATTC	1467
Qy	477	ProValAlaIleLeuSerLysGluArgAlaMetLeuGluAsnIleMetArgAlaLys	496
Ds	1468	AAACCTATCAACGGAATTTAACAAACAACGTCACAAATTAATGATCAACTTTCAGTGTG	1527
Qy	497	ValGlyLeuAlaProGluAsnMetIleLeuGlu	508
Ds	1528	GTTGGTTTACCAATGTAACAACGAATTAAGAATTGAA	1563

Sequence 12608, Application US/09949016
Patent No. 5812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 60/231,498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 12608
LENGTH: 77626
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(77626)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12608

Alignment Scores:
Pred. No.: 1,796-141 Length: 77626
Score: 1229.50 Matches: 313
Percent Similarity: 41.22% Conservative: 86
Best Local Similarity: 32.33% Mismatches: 106
Query Match: 46.73% Indels: 466
DB: Gaps: 12
US-10-718-952-12 (1-510) x US-09-949-016-12608 (1-77626)

QY 6 PheHysValAlGusePProAsnValLyStryrThrGluThrGluIleGlnSerValTyzAsn 25
DB 32903 TTCTTCGTCGAGAGCCCGGACCTGCTACCGCCCGAGCCATCGAGCGCAATATACAG 32962
QY 26 TygIuThrThrGluLeuValHisGlu----- 34
DB 32963 TACCGAGACGACCGCGCTACGCCCGAGGGTGGCTTCTCAAGCTGACCCAGGGGCTGAG 33022
QY 35 -----AsnArgAsnGlyThrTyrlnTrpIle 43
DB 33023 GAAGGAGGCTGCTGGGAGAGCGGGGCGCTCGGGAGAAAGGACCGGAGATT 33082
QY 44 -----VallyPProLySserValAsnTyrlnPhelysThrsAn 56
DB 33083 CACTGCGCGCTGCTTCAGGTGCACCCACGTCACGCGCTTCACCTTCGAGACCGCC 33142
QY 57 ThrHisValPProLyLeuGlyValMetLeuValGlyTrpGlyValAsnAsnGlySerThr 76
DB 33143 CGGACGAGTCCCGGCTCGGGGTCAAGCTGTGCGTGGGGCGGAAACAACGCTCCACA 33202
QY 77 LeuThrGlyGlyValIleIleAsnArgGluAspIleSerTrpAlaThrLyAsp----- 94
DB 33203 CTACACGCGCGCGCTGCGCATGCACTGCGCTTGTCTGCGCCACGCGGACGCGCC 33262
QY 94 ----- 94
DB 33263 AAGGTGGGGGCGGAGGGGCGTGTGAGTGAAGAGCTTGTGAGATCTTGGCGGCG 33322
QY 94 ----- 94
DB 33323 GGGCGGAGCTCTGACAGGGGGCGGGCTCTGCGGACCAAGGGGGCTGTAGGGGGTGGGCT 33382
QY 94 ----- 94
DB 33383 TCCGCGGGGCGAGGGGCTGCAAGGGGCGAAGGACCTGCCGGGTCTGAGAGCTGTAGTGG 33442

QY 94 ----- 94
DB 33443 GCGAGGACCTGACAGGGGCGAAGCTCTGTGGGCGGAGAGGCTTGCAGGGGGCGGGCC 33502
QY 94 ----- 94
DB 33503 TGGCTGACAGGGGACCCCGGTGAAGGTTTGGCGAGGTAGCGGGGACTCTCGGAGTCT 33562
QY 94 ----- 94
DB 33563 TACAGGGCGAGCTTAAGTGTCCGAGAGTCTCAAGGAGGGGCTTGGGGGCGAGCGCTT 33622
QY 95 -----LysIle----- 96
DB 33623 CCGTCTGCTGTGAGCTTGAATGAATAATGCCGAAAGTGGGGCGGGCTTGAAGT 33682
QY 96 ----- 96
DB 33683 CTAGGTGGGGGCGGAGCGGGGCTGGGCTTTCACCACTCCGACCTGCGCGT 33742
QY 97 -----GlnGlnIleAsnTyrlPheGlySerLeuThrGlnIleAsnAlaIleAr 112
DB 33743 GTCTGCTGACCCCGAGAGGCGCAACTACTAGGCTGCTGACTCAGCGGGGACCGTGA 33802
QY 112 gValGly---SerPheGlnGlyGluIleTyrlaIleProPheLySserLeuLeuPhe 131
DB 33803 CCGGGCTTGAACCGCGAGGCGCAGAGGTGTCTGACCTTTCAGCGGGTCTGCCAT 33862
QY 131 tValAsnProAspAspIleValPheGly----- 140
DB 33863 GGTGGCGCCCAACGACTCTGTTTGA--TGTGGGCGGAGCCCTGGGCGGGGTGGGGCG 33921
QY 140 ----- 140
DB 33922 GGATGGAGATGGGCTGTGAGAGGCGCCAGATCCGGGCGGAGGTGAGGATCC 33981
QY 141 -----GlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaIleAlaValPh 159
DB 33982 CCGAGGCTGGGACATCTGTGCTGATACCTGACCGACGAGCGCGCGCAAGGTGT 34041
QY 159 eAspIleAspLeuGlnLySglInLeuArgProTyrlMetGluSerMetValProLeuProG 179
DB 34042 GAGCTGGGGGCTGCGAGAGCAACTGTGGCCGCACTGAGAGCCCTCGGGCCCTTC 34101
QY 179 yLleTyrlAspProAspPheIleAlaIleAsnGlnGluIleArgAlaAsnAsnValIleLy 199
DB 34102 TGTTCACATCCCGAATTCACTGCGGCGCAACAGAGCGCGCGGACAACCTCATCCC 34161
QY 199 sGlyThrlYsGlnGlnGlnVal----- 206
DB 34162 AGGCTCGGTGCGCAGAGGTGTGTCTCCATCCGCACTCCTTGTCTTCCACTGT 34221
QY 206 ----- 206
DB 34222 CTTCGCACTTGTTCTCTCTTTTGGCCCAACCGCGCAATCCAGGCTCACTTACAC 34281
QY 206 ----- 206
DB 34282 CGGGAACTGAGTAGCCAGGCTCTGTGAGTCCGAGATGTATCAACAACATTCCCTA 34341
QY 207 -----GlnGlnIleIleIle 211
DB 34342 CACCTCTCACTCTCCACACCTCACTCAATCCAAACCAAGCTGAGAGCAATCCGAG 34401
QY 211 aAspIleLyValAlaPheLySglValaIleThrlYsValAspLyValIleValLeuTrpThrAl 231
DB 34402 GGAATCCGAGACTTCCGCTTACGCGCGGGCTGACAAAGTCAATAGTGTGTGAGCGC 34461
QY 231 aAsnThrlGluArgTyrlSerAsnLeuValIleGlyLeuAsnAspThrlMetGluAsnLeu 251
DB 34462 GAACACGAGGCGCTTCTGTGAGGATTCAGGCTCAACACACACAGCGGAACCTGCT 34521
QY 251 uAlaAlaValAspArgAsnGlu-Ala----- 259

```
Db 34522 GGCACCAATTAGGTGGGGGAAAGCCCGGGTGAAGGGGTCAAGCAGAGGCCAGTCAACA 34581
Qy 260 ----- -GluIleSerProSert 265
Db 34582 CACAGCTGTGTGTACTACTGCCCCCTTGGCCACAGCTCGGTGTGAAGGTGTGCCCCCTTCA 34641
Qy 265 hIleuTYrAlAlIeAlaCyValMetGluAsnValProPheIleAsnGlySerProGlnA 285
Db 34642 CGCTCTTGCGCGTGGCAGACATCTGGAAGGGCTGTGCTTCTCATGTGGTCTCCGACAGA 34701
Qy 285 snThrPheValProGlyLeuIleAspLeuAlAlaArgAsnThrLeuIleGlyIYA 305
Db 34702 AACCCCTGGTGCCTGGAGCTGTGAGCTCGGTGCAGACACCGGGTCTTTGTGGGGCGAG 34761
Qy 305 sPAspPheIleSerGlyInThrIleMetIleSerValIleuValAspPheLeuValGlyA 325
Db 34762 ATGACTTCAGAGTCAGGCCAGACCAAGTCAAGTCCGTGCTTGAGCTTCTCATTTGGCT 34821
Qy 325 IaGlyIleIys----- 328
Db 34822 CCGGGCTCAAGGTGGGTGGGCTTAGAGAGCTCCGAGTGCAGAGAAAGGGGCTTGGGCGCG 34881
Qy 328 ----- 328
Db 34882 GAGCACTGGCTTGTGGGGCCGAGGGCCTGAGCTGTGGGCTTCCCTTGTACCCACA 34941
Qy 329 --ProThSerIleValSerIleYrAsnHisIleuGlyAsnAsnAspGlyMetAsnLeuSerA 348
Db 34942 GACCAATG-TCATACGAGTTCACACCACTGGGCAACAACATGAGGAGAACTATATGCG 35000
Qy 348 IaProGlnThrPheArgSerIleYrGluIleSerIleSerAsnValValAspAspMetValA 368
Db 35001 CGCCATTGAGATTCGGCTCTAAGAGAGGTGTCCAGAGCAACCTGTGTGACGACATGGTGC 35060
Qy 368 snSerAsnAlAlaIleuTYrGluPProGlyGluHisProAspHisVal----- 383
Db 35061 AGAGCAACCCAGTGTCTATACGCCCGGAGAGAGCTGACCACTG-CGTGGTGGGGCG 35119
Qy 383 ----- 383
Db 35120 CGGGCGCGGGCGGGGCTTCCCGGAGAGGGGTGGCAGAGGCCCGGCAAACTCCTGCTG 35179
Qy 384 -----ValValIleYrTYrValProTYrValGlyAspSerIleYrAsnIleAsnAspG 401
Db 35180 CACTCCAGGTGTCAATCAAGTATGTGCGGTACGTGGGTGACAGCAAGCGCGGTGANTG 35239
Qy 401 luTYrThSerGluIlePheMetGlyIYrAsnThrIleValIleuHisAsnThrCySG 421
Db 35240 AGTATACCTCGGAGCTGATGTGGCGGGAACCAACACACTGTGCTGCACAACAGTGTG 35289
Qy 421 lu----- 421
Db 35300 AGGTGCGGGCGGCTTACAGGGGCTGTATGGAGGCGAGGGCTTGGCCACCGCCCCACTCT 35359
Qy 422 -----AspSerIleuLeuAlAlaProIleIleLeuAspLeuVal 434
Db 35360 CCTGACCAACCGCCCGGAGAGACTCGCTGCTGGCGGACCATATATGCTGGAACCTTACG 35419
Qy 435 IeuleuAlaGluLeuSerThrArgIleGlnPheIYrAlaGluAsnGluGlyIYrPheHis 454
Db 35420 CTGCTGACCGAGCTGTGCAGCGCTGAGCTTCTGACATGACATGCCGAGCCGAGCGCAG 35479
Qy 455 SerPheHisProValAlaThrIleIleuSerTYrLeuThrIYrAlaProLeuValProPro 474
Db 35480 ACCTTCCACCCGCTGTCTCTGCTCAAGCTTCTTCAAGGCCCACTAGTGGCGGCC 35539
Qy 475 GlyThrProValValAsnAlaLeuSerIYrGlnArgAlaMetLeuGluAsnIle----- 492
Db 35540 GGCAGCCCGGTGTCAATGCGCTTTTCCGACGAGCCAGCTGCATGAGAAACATCTCTCAGG 35589
Qy 492 ----- 492
```

```
Db 35600 TGACCCCAAGCCTCCAGATTCCCGCCCGGGCCAGATACCCCTGTGCTAAGCTAAGCA 35659
Qy 492 ----- 492
Db 35660 CAGGCTTGGGGGCTGTAGGTGGCAGGATACCCCTTTCTTCCCTGTGGGTGTCTGG 35719
Qy 493 -----Met-ArgAlaCyValGlyIleuAlaProGlnAsnAsnMe 505
Db 35720 CTGTGCTAAGCTGTGTGACTCTCCAGGGCTGTGGGTGGCTCCCGCACAGAACCAT 35779
Qy 505 tIleuGlnTYrIYs 510
Db 35780 GCTCTGGAAACAAA 35795

RESULT 12
US-09-397-787-34
/ Sequence 34, Application US/09397787
/ Patent No. 6468758
/ GENERAL INFORMATION:
/ APPLICANT: Benson, Darin R.
/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: King, Gordon E.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
/ FILE REFERENCE: 210121.466C2
/ CURRENT FILING DATE: 1999-09-16
/ NUMBER OF SEQ ID NOS: 334
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 34
/ LENGTH: 1231
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-09-397-787-34

Alignment Scores:
Pred. No.: 1,08e-128 Length: 1231
Score: 1103.50 Matches: 210
Percent Similarity: 81.29% Conservative: 55
Best Local Similarity: 64.42% Mismatches: 58
Query Match: 41.94% Indels: 3
DB: Gaps: 1

US-10-718-952-12 (1-510) x US-09-397-787-34 (1-1231)
Qy 187 AlAlaAsnGlnGluGluArgAla-AsnAsnValIleIYsGlyThrIYsGlnGlnIYnVa 206
Db 5 GCGGCACCAACGAGCGCGCGCGGCGGCAACCTCATCCAGGCTCGGTGCGCAGCAGCT 64
Qy 206 lGlnGlnIleIleIYsAspIleYsAlaPheIYsGluAlaThrIYsValAspIYsValVa 226
Db 65 GGAGAGATCCGCGAGGAGCATCCGAGACTTCCGTCTAAGCGGGCGGTGGACAAAGTCAAT 124
Qy 226 lValIleuTYrThAlaAsnThrGluArgTYrSerAsnIleuValValGlyLeuAsnAspTh 246
Db 125 AGTCTGTGACGCGGCAACGAGAGCCCTTCTGTGAGTGTATTCAGAGCCTCAACGACAC 184
Qy 246 rMetGluAsnIleuAlaAlaValAspA-ArgAsnGluAlaGluIleSerProSerThIle 266
Db 185 AGCCGAGAAACCTGTGCGCACCATYTAGCTCGGT---CTGAGAGGTGTGCGCTTCCAGCT 241
Qy 266 uTYrAlAlaIleAlaCyValMetGluAsnValProPheIleAsnGlySerProGlnAsnTh 286
Db 242 CTGTGCGGTGCGCAGATCTGAGGGCTGTGCTTCTCAATGAGTCTCCGCAAGAACAC 301
Qy 286 rPheValProGlyLeuIleAspLeuAlAlaIleAlaArgAsnThrIleuIleGlyIYsAspAs 306
Db 302 CCGTGTGCGCGAGACTTGTAGCTCGGTGCGCAGACCGGGTCTTTGTGGCGGAGATGA 361
Qy 306 pPheIYsSerGlyGlnThrIYsMetIYsSerValIleuValAspPheLeuValGlyAlaG 326
Db 362 CTTCAGTCAAGGCGAGACCAAAATCAAGTCAAGTCCGTGCTTGTGAACTTCTCATTTGGCTCGG 421
```

QY 326 YLELYSProthrSerileValSerTyrrhAnstleuGlyAaAnaAnaArgLysMetAsn-L 346
Db 422 CTTAAAGACATGTCATGATGATACCAACCTGGGCAACAGATGGGAGAACCC 481
QY 346 euserAlaProGlnThrPheArgSerLysGluileSerLysSerAnaValAlaAspPM 366
Db 482 TATGGGCCCATTCGACGTTCCGCTCTAAGAGAGTGTCCAAAGCAACGCGTGGAGACA 541
QY 366 etValAsnSerAnaAlaileuTyrrGluProGlyGluHlAProAphlValValI 386
Db 542 TGGTCAGAGCAACCAAGCTCTATACCCCGCGAGAGAGCTGACCTCGGTGTCA 601
QY 366 lelyTyrrValProTyrrValGlyAAspSerLysArgAlaMetAspGluTyrrThrSerGlu 406
Db 602 TCAAGTATGTCGCTGATCGGGGTGACAGACGCGGCTGATGATATACCTGGAGAC 661
QY 406 lPheMetGlyGlyLysAnthrIleValLeuHlAsnThrCysGluAAspSerLeuLeu 426
Db 662 TGATGCTGGGCGGAACAACACACTGTGCTGCACAAACGTGTAGAGACTGCTGCTGG 721
QY 426 lAlaProIlelleuAAspLeuValLeuLeuAlaGluLeuSerThrArgIleGlnPheL 446
Db 722 CCGCACCCATCATGCTGGACCTAGCGCTGCTGACCGAGCTGTGCGACGCGTGAAGCTTCT 781
QY 446 yValAGluAAsnGluGlyLysPheHlAspSerPheHlAspProValAlaThrIleLeuSerTyrrL 466
Db 782 GCACGTGACATGACCCCGGACGCGGACGCTTCCACCCGCTGTGCTGCTGCTGCTGCTTCC 841
QY 466 eutThrLysAlaProLeuValIProProGlyLysThrProValAlaAnaAlaLeuSerLysGlnA 486
Db 842 TCTTCAAGGCGGCACTAGTGGCGGCGGCGGAGCGCGGTGTCAATGGCTTTTCCGCGAC 901
QY 486 rGAlaMetLeuGluAAsnIleMetArgAlaCysValGlyLeuAlaProGluAAsnMetI 506
Db 902 GCAGCTGATCGAAGAACATCTCTAGGGGCTGTGGTGGGCTCCCGCACAGAACACATGCTC 961
QY 506 lelyGluTyrrLys 510
Db 962 TCTGTGAACACAA 975

RESULT 13

US-09-902-540-1257/C
; Sequence 1257, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1257
; LENGTH: 34316
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1257

Alignment Scores:

Score: 1.02e-53 Length: 34316
Percent Similarity: 529.00 Matches: 141
Best Local Similarity: 46.92% Conservative: 103
Query Match: 27.12% Mismatches: 206
Query Match: 20.11% Indels: 71
Gaps: 13

US-10-718-952-12 (1-510) x US-09-902-540-1257 (1-34316)

QY 17 GluThrGluileGlnSerValTyrrAsnTyrrGluThrGluLeu-----Val 32
Db 20023 GAAGTTGAAGTGAAGACGTTTAGATATCAACCATCCGCTTTGGAGATTTTCCGCATG 19964
QY 33 HieGlu-----AsnArgAnGlyThr 39
Db 19963 CACGACCGGATTCGTAACGACCCCGCATACGCGCCGATGAGTATAGCTTAAAGAGAACCA 19904
QY 40 TyrrGlnThrIleValLysProLysSerValAsnTyrrGlnPheLysThrAsnThrIleVal 59
Db 19903 TCACGATGAAACAAAGAGTGTGCGGAAGCCGAGCGGACAGCTCCGCTCGG-ATY 19845
QY 60 ProLysLeuGlyValMetLeuValGlyTyrrGlyLysAnaAnGlySerThrLeuThrGly 79
Db 19844 CCGGCTGGGCGCTGTC-----TCCACCAAGCTGATGGCG 19809
QY 80 GlyValIleAlaAsnArgGluAAspIleSerTrpAlaThrLysAAspLysIleGlnGlnAla 99
Db 19808 GGTGTGAGCTGGCGCGGACAGGCG-----AAGGCT 19779
QY 100 AsnTyrrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySer---PheGlnGly 118
Db 19778 CACCCCATTTGGGTGCTGCTACCGCAGATGGGACCGCCCGCTGGGAGACGAGCCGACGG 19719
QY 119 GluGluIleTyrrAlaProPheLysSerLeuLeuProMetValAsnProAAspIleVal 138
Db 19718 CGCACCGTG-----AACCTCAACGAGCTGGTGGCTCTGGGCGAGCTGAAGACGTTGCC 19665
QY 139 PheGlyGlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysVal 158
Db 19664 TTGCGCGCTGGACATCATCGTGTAGAGACCGCATACGAAGTGGCGCTGGCGCGGTG 19605
QY 159 PheAspIleAAspLeuGluLysGlnLeuArgProTyrrMetGluSerMetValIProLeuPro 178
Db 19604 CTCAGCGCAAGACACTGTAAGAGGTGAAGCGCTTCCCTCCAGAGCATCAAGCCGAAGAG 19545
QY 179 GlyIleTyrrAspProAAspPheIleAlaIleAnGlnGluGluArgAlaAsnAnValIle 198
Db 19544 GCGGTGACAGACCTGAGTGTGCTGGCGCGCATTCAGAGCC-----AACCAATTC 19497
QY 199 LysGlyThrLys-----GlnGluGlnValGlnGlnIleIleLysAAspIleLysAlaPhe 216
Db 19496 AAGGCCACCAAGACCCACCGGAGAGATCGAAGCGCTGGCCGACGATCCGCGCATTC 19437
QY 217 LysGluAlaThrLysValAAspLysValValIleValLeuTyrrPheAlaAsnThrGluArgTyrr 236
Db 19436 AAGAAAGAGCTCAACGCGACCGCGCGTATGGGTGTGTCAGACGCTGAGACGCTTC 19377
QY 237 SerAsnLeuValIleGlyLeuAAspThrMetGluAnLeuLeuAlaIleValAAspArg 256
Db 19376 GGTCCGCTGCCCGAGTCTTCAAG--ACGCTGGCGGCTTTCGAGAGCGCTGAGCAG 19320
QY 257 AsnGluAlaGluileSerProSerThrLeuTyrrAlaIleAlaCysValMetGluAAsnVal 276
Db 19319 AAGAGCCCGGACATCAACCCACGCGGTGTATACCTAGCGGCGCATCAAGAGGCGGTG 19260
QY 277 ProPheIleAnGlySerProGln--AsnThrPheValIProGlyLeuIleAAspLeuAla 295
Db 19259 CCTTTCGGAAGCGCACCGCCCAACGCGCTGAGACGCGCGGCTTCAGAGAGTGGCC 19200
QY 296 IleAlaArgAnThrIleIleGlyLysAAspPheLysSerGlyGlnThrLysMetLys 315
Db 19199 AAGCAGAGTGGTGGCGGCTGCGCGCGCATCAAGACGCGCATGATGTAG 19140
QY 316 SerValLeuValAAspPheLeuValGlyValAGlyIleLysProThrSerIleValSerTyrr 335
Db 19139 ACGGTGATCGGCGCGCGCTTAAGCGCCGATGCTGTGATGCTGCTTCCAC 19080
QY 336 AsnHlAsnGluLysAnaAnaAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLys 355
Db 19079 AACATCTGGGCAACCGGACGCGGAGGTGCTGAGACGACCCCGCGCTTCAAGGCGCAAG 19020
QY 356 GluIleSerLysSerAnaValAlaAAspAspMetValAsnSerAnaAlaIleLeuTyrrGlu 375

```
Db 19019 GAAGTCACCAAGTCGAGTCGTGACACATCTG-----CAG 18981
Qy 376 ProgluyluHleuProAap-----HisValValIleuetyrVal 389
Db 18980 CCCGACCTGTACCCGGACCTGTACAGAAGTACGGCGCAAGGTGGCCATCCACTACTAC 18921
Qy 390 ProTyrValIGlyAspSerlyeAraGlaMetAspGluTyrThrSerGluIlePheMetGly 409
Db 18920 CCGCCCCCGCGGACGCCAAGGAGGTTGGGACCAACATCGACATCCGAGTGGCTGGC 18861
Qy 410 GlyLysAsnThrIleValIleuHisAsnThrCysGluAspSerLeuLeuAlaProIle 429
Db 18860 TATCCGATGACGATCAAGGTCAACTCTCTGCGCCGACCTCATCTGCGCGCCGCTG 18801
Qy 430 IleuAspLeuValIleuLeuAlaGluLeuSerThrArgGlnPheLysAlaGluAsn 449
Db 18800 GTGCTGACATCGCCCTGTCTTGGACCTGGCCAAAGCGAGTGGCGGCGCATCCAG 18741
Qy 450 GluGlyLysPheHisSerPheHisProValAlaThrIleLeuSerTyrLeuThrLysAla 469
Db 18740 GAGTGG-----ATGCTCTTCTACTTCAAGACC 18714
Qy 470 ProleuValProProGlyThrProValValAsnAlaLeuSerLysGlnArgAlaMetLeu 489
Db 18713 CCCATGGCGCAGCGCGGTCTGCGCGTGGAGCAGCAGCTTTCATCCAGTCAACGAGCTG 18654
Qy 490 GluAsnIleMetArgAlaCysValIGlyLeuAlaProGluAsnAsnMetIleuGluTyr 509
Db 18653 AAGAACACGCTCGCGGTCTGCGGGCGAGAGACCCCATCACCTCGGGCTCGACTAT 18594

RESULT 14
US-09-902-540-5256 Application US/09902540
; Sequence 5256, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5256
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-5256

Alignment Scores:
Pred. No.: 4,936-56 Length: 1335
Score: 528.00 Matches: 135
Percent Similarity: 49.24% Conservative: 91
Best Local Similarity: 29.41% Mismatches: 187
Query Match: 20.07% Indels: 46
DB: 4 Gaps: 11

US-10-718-952-12 (1-510) x US-09-902-540-5256 (1-1335)
Qy 61 LysLeuGluValMetLeuValGlyTrrGlyGlyAsnAsnGlySerThrLeuThrGlyGly 80
Db 37 AAGCTCGCGGTCTTAATTCGGGCTCTGGCGCTGTGTC--ACCACTGATGGCGGGGT 93
Qy 81 ValIleAlaAsnArgGluAspIleSerTrpAlaThrLysAspLysIleGlnGlnAlaAsn 100
Db 94 GTCGAGCTCGCGCGGCAAGGC-----AAGGTCAAC 123
Qy 101 TyrPheGlySerLeuThrGlnAlaSerAlaIleArgValGlySer--PheGlnGlyGlu 119
|||||
```

```
Db 124 CCATTTGGGTCTCACGACAGATGAGGACCGCCGCTGGGGAAGCGGACCGGCGGC 183
Qy 120 GluIleTyrAlaProPheLysSerLeuProMetValAsnProAspIleValPhe 139
Db 184 ACCGTG-----AAGCTCAACGACCTGTGCTTGTGGCGAGCTGGAAGAGCTTGCTTC 237
Qy 140 GlyIlyTrpAspIleSerAsnMetAsnLeuAlaAspAlaMetAlaArgAlaLysValPhe 159
Db 238 GCGCCTGTGGACATCATCTCGTAGAGGACGACATACGAAGTGGCGCTGCGGCTGGCTC 297
Qy 160 AspIleAspLeuGlnLysGlnLeuArgProTyrMetGluSerMetValProLeuProGly 179
Db 298 AGCCACAAAGCACTGGAGAGGTAAAGCCGTTCTCCAGACATCAAGCCGACAGGCG 357
Qy 180 IleTyrAspProAspPheIleAlaAlaAsnGlnGluArgAlaAsnAsnValIleLys 199
Db 358 GTGCACGACCTTGAGTTCTGTGCGCGCATCGAGGC-----AACCATCAAG 405
Qy 200 GlyThrLys-----GlnGlnGlnValGlnGlnIleIleLysAspIleLysAlaPheLys 217
Db 406 GCCACAAAGACGACCGCGAGAGCATGACGAAGCGCTGGCGCGACATCCGCACTTCAAG 465
Qy 218 GluAlaThrLysValAspLysValValValLeuThrPheAlaAsnThrGluArgTyrSer 237
Db 466 AAGAGCTCAACGGCAGCGCGCGCTGATGTGTGACAGCAGCTTGACCTCCGT 525
Qy 238 AsnLeuValValIGlyLeuAsnAspThrMetGluAsnLeuAlaAlaValAspArgAsn 257
Db 526 CCGCTGCCCGAGTCTTCAAG--ACGCTGCGCGCTTCGAGAGGCGCTGAGAGAGAC 582
Qy 258 GluAlaGluIleSerProSerThrLeuTyrAlaIleAlaCysValMetGluAsnValPro 277
Db 583 AGCCCGACATCAACCCACCGCGCTGTACCTACGCGGCATCAAGAGAGGCGTGGCC 642
Qy 278 PheIleAsnGlySerProGln--AsnThrPheValProGlyLeuIleAspLeuAlaIle 296
Db 643 TTTCGGAAGCGCCGCCAACGCCAGCGGTGACACGCGCGCTCCAGAGATGACCAAG 702
Qy 297 AlaArgAsnThrIleuIleGlyIGlyAspAspPheLysSerGlyGlnThrLysMetLysSer 316
Db 703 CAGAGATCGGTGGCGCGTCCGCGCGGACCTCAAGCGCCGACGATGATGAAGACG 762
Qy 317 ValLeuValAspPheLeuValIGlyAlaGlyIleLysProThrSerIleValSerTyrAsn 336
Db 763 GTCATCGCGCGCGGCTCAAGCGCCGATGCTGTGATGAGCTGATGCTTCCACCAAC 822
Qy 337 HisLeuGlyAsnAsnAspGlyMetAsnLeuSerAlaProGlnThrPheArgSerLysGlu 356
Db 823 ATCTGGGCAACCGCGAGCGGAGGTCTGGACGACCCCGGCTTCAAGGCCAAGGAA 882
Qy 357 IleSerLysSerAsnValValAspAspMetValAsnSerAsnAlaIleLeuTyrGluPro 376
Db 883 GTCAACAAAGTCAAGCGTCTGGACACCACTCTG-----CAGCC 921
Qy 377 GlyLysHisProAsp-----HisValValIleLysTyrValPro 390
Db 922 GACCTGTACCGGACGTGTACAGAAGTACGCGCAAGGTGGCCATCCACTATACCGG 981
Qy 391 TyrValIGlyAspSerLysArgAlaMetAspGluTyrThrSerGluIlePheMetGlyGly 410
Db 982 CCGCGCGCGGACGCGAGAGGAGGTTGGGACAACTCGACATCACCGGATGGCTGGCAT 1041
Qy 411 LysAsnThrIleValIleuHisAsnThrCysGluAspSerLeuLeuAlaProIleIle 430
Db 1042 CCGATGACGATCAAGGTCAACTTCTCTGCGCGACATTCATCTTGGCGCGCGCTGTG 1101
Qy 431 LeuAspLeuValLeuLeuAlaGluLeuSerThrArgIleGlnPheLysAlaGluAsnGlu 450
Db 1102 CTGACATCGCGCTTCTCTGACCTGCGCAAGCGGCTGAGTGGCGGCGCATCAGAG 1161
Qy 451 GlyLysPheHisSerPheHisProValAlaThrIleLeuSerTyrLeuThrLysAlaPro 470
Db 1162 TGG-----ATGCTCTTCTACTTCAAGACCCCC 1188
```

QY 471 LeuValProProGlyThrProValAlaAlaLeuSerLysGluArgAlaMetLeuGlu 490
Db 1189 ATGGCGCCAGCCGGGTCTGCGGTGAGACGACCTTTCATCTCAGCTACCAAGCTGAAG 1248
QY 491 AsnIleMetArgAlaCysValGlyLeuAlaProGluAsnAsnMetIleLeuGluTyr 509
Db 1249 AACACGCTGCGGTGCTGCGGGGAGACCCCATCACCACCTCGGGCTGACTAT 1305

RESULT 15

US-09-313-294A-4684
; Sequence 4684, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 4684
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700348858H1
; NAME/KEY: unsure
; LOCATION: 185, 272, 275, 281, 283, 288
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4684

Alignment Scores:

Pred. No.:	9,51e-11	Length:	294
Score:	160.00	Matches:	28
Percent Similarity:	82.61%	Conservative:	10
Best Local Similarity:	60.87%	Mismatches:	8
Query Match:	6.08%	Indels:	0
DB:	4	Gaps:	0

US-10-718-952-12 (1-510) x US-09-313-294A-4684 (1-294)

QY 1 MetPheIleGluAsnPhelyValGluSerProAsnValLysTyrThrGluThrGluIle 20
Db 125 ATGTCATCGAGAGCTTCGCGTCCGAGAGCCCCACGTCGGTACGGCCCGACGAGATC 184
QY 21 GlnSerValTyrAsnTyrGluThrThrGluLeuValHisGluAsnArgAsnGlyThrTyr 40
Db 185 NTCTCGAGGTACCGGTACGACGAGCGAGCTGTGCAAGAGCCAAAGACGGCGCTCC 244
QY 41 GlnTPrIleValLysPro 46
Db 245 CGCTGGGTCTCGGCC 262

Search completed: June 7, 2005, 22:59:18
Job time : 307 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:49:11 ; Search time 7680.29 Seconds
(without alignments)
11103.899 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760

Sequence: 1 cctctctctctctctcttctgt.....aatgtagtataatttctgt 1760

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1760	100.0	1782	6	BD075266	BD075266 Soybean p
2	1777	97.6	1739	6	AY038802	AY038802 Glycine m
3	1482	84.2	1533	6	BD075269	BD075269 Soybean p
4	1454	82.6	1791	8	AF293970	AF293970 Glycine m
5	65	3.7	62713	8	AP006418	AP006418 Lotus cor
6	59	3.4	2994	8	AF282263	AF282263 Phaseolus
7	45	2.6	1391	8	AY005128	AY005128 Actinidia
8	40	2.3	119858	8	AP006419	AP006419 Lotus cor
9	35	2.0	1978	8	CPINOIG	U32632 C.pardisi
10	32	1.8	2053	8	MCU32511	U32511 Mesembryant
11	32	1.8	143012	8	AC125389	AC125389 Medicago
12	26	1.5	1533	8	AY323824	AY323824 Xerophyta
13	26	1.5	1564	8	AY096554	AY096554 Arabidops
14	26	1.5	1781	6	AX054630	AX054630 Sequence
15	26	1.5	1781	8	BN06307	BN06307 Brasica na
16	26	1.5	1845	8	AY028259	AY028259 Avicennia
17	26	1.5	1863	8	AY065415	AY065415 Arabidops
18	26	1.5	1986	8	AF433879	AF433879 Suaeda ma
19	26	1.5	80088	8	AT3131P6	AT356332 Arabidops

20	24	1.4	35	6	BD075267	BD075267 Soybean p
21	21	1.4	149633	2	AC087428	AC087428 Homo sapi
22	24	1.4	187003	9	AC023353	AC023353 Homo sapi
23	23	1.3	1535	6	BD073472	BD073472 Regulated
24	23	1.3	1538	8	AF120146	AF120146 Trilicium
25	23	1.3	1538	8	AF120147	AF120147 Trilicium
26	23	1.3	1538	8	AF120148	AF120148 Trilicium
27	23	1.3	1539	8	AF412340	AF412340 Porteresi
28	23	1.3	1578	6	E13612	E13612 gDNA encodi
29	23	1.3	1578	6	E17392	E17392 gDNA encodi
30	23	1.3	1665	6	BD073470	BD073470 Regulated
31	23	1.3	1665	8	AF056326	AF056326 Zea mays
32	23	1.3	1868	8	AB012107	AB012107 Oryza sat
33	23	1.3	1914	8	AK013501	AK013501 Oryza sat
34	23	1.3	1915	8	AK058750	AK058750 Oryza sat
35	23	1.3	1931	6	AR137808	AR137808 Sequence
36	23	1.3	1931	6	AR170167	AR170167 Sequence
37	23	1.3	1931	6	AB032073	AB032073 Nicotiana
38	23	1.3	1936	8	AB059557	AB059557 Avena sat
39	23	1.3	1950	6	E27176	E27176 Novel INPS
40	23	1.3	1959	8	AB009881	AB009881 Nicotiana
41	23	1.3	1959	8	BT013505	BT013505 Lycopersi
42	23	1.3	3546	6	AR137811	AR137811 Sequence
43	23	1.3	3546	6	AR137812	AR137812 Sequence
44	23	1.3	3546	6	AR170170	AR170170 Sequence
45	23	1.3	3546	6	AR170171	AR170171 Sequence

ALIGNMENTS

RESULT 1	BD075266	1782 bp	DNA	linear	PAT 27-AVG-2002
LOCUS	BD075266	Soybean plant producing seeds with reduced levels of raffinose			
DEFINITION	BD075266	saccharides and phytic acid.			
ACCESSION	BD075266.1	GI:22620869			
VERSION	BD075266	JP 2001519665-A/1.			
KEYWORDS	JP 2001519665-A/1.	unidentified			
SOURCE	unidentified	unclassified			
ORGANISM	unclassified.	unclassified.			
REFERENCE	1 (bases 1 to 1782)				
AUTHORS	Hitz,W.D. and Sebastian,S.A.				
TITLE	Soybean plant producing seeds with reduced levels of raffinose				
JOURNAL	patent: JP 2001519665-A 1 23-OCT-2001;				
COMMENT	RI DU POINT DE NEMOURS AND CO				
	OS Soybean line LR13				
	PN JP 2001519665-A/1				
	PD 23-OCT-2001				
	PP 07-APR-1998 JP 1998543012				
	PR 08-APR-1997 US 08/835751				
	PI WILLIAM DEAN HITZ,SCOTT ANTHONY SEBASTIAN				
	PC C12N15/52,C12N15/62,C12N15/11,C12N9/90,A01H5/00 CC				
	Strandedness: Double;				
	CC Topology: Linear;				
	CC Soybean plant producing seeds with reduced				
	levels of raffinose				
	CC and phytic acid				
	PH Key Location/Qualifiers				
	FT CDS 54..1586.				
FEATURES	Location/Qualifiers				
source	1..1782				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
ORIGIN					
Query Match	100.0%;	Score 1760;	DB 6;	Length 1782;	
Best Local Similarity	100.0%;	Pred.No. 0;			
Matches 1760;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	

```

QY      1 CTCTCTTATTCCTTTGTATTTTCATTCATTTCTTAATCTTTGTGAAAAAATATGTTCA 60
Db      1 CTCTCTTATTCCTTTGTATTTTCATTCATTTCTTAATCTTTGTGAAAAAATATGTTCA 60
QY      61 TCGAATAATTTAAGGTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATTGAGTCGG 120
Db      61 TCGAATAATTTAAGGTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATTGAGTCGG 120
QY      121 TGTACACTACGAAAACCAACGAACTGTTCACAGAGAACGGAATGGCACCTATCATGTGA 180
Db      121 TGTACACTACGAAAACCAACGAACTGTTCACAGAGAACGGAATGGCACCTATCATGTGA 180
QY      181 TTGTCAAAACCAATCTGTCAAAATCGAATTTAAAACCAACCAATCTGTCTTAATTAG 240
Db      181 TTGTCAAAACCAATCTGTCAAAATCGAATTTAAAACCAACCAATCTGTCTTAATTAG 240
QY      241 GGGTAATGCTTGTGGGTGGGGTGGAAACAAAGGCTCAACCGTCAACGGTGGTGTATTG 300
Db      241 GGGTAATGCTTGTGGGTGGGGTGGAAACAAAGGCTCAACCGTCAACGGTGGTGTATTG 300
QY      301 CTAAACGAGAGGACATTTATGAGGCTTAAAGGACAAAGTCAACAAAGCAATTAATTG 360
Db      301 CTAAACGAGAGGACATTTATGAGGCTTAAAGGACAAAGTCAACAAAGCAATTAATTG 360
QY      361 GGTCCCTCAACCAAGCTGAGCTATCCGAGTTGGGTCTTCCAGGGAGAGAAATCTATG 420
Db      361 GGTCCCTCAACCAAGCTGAGCTATCCGAGTTGGGTCTTCCAGGGAGAGAAATCTATG 420
QY      421 CCCCATTCAGAGACCTGTTCCAAATGTTAACCTGACGACATTTGTGTTGGGGAGTGG 480
Db      421 CCCCATTCAGAGACCTGTTCCAAATGTTAACCTGACGACATTTGTGTTGGGGAGTGG 480
QY      481 AATATGACAAATGAACCTGTGCTGATGCCATGCGGCGAAAGGTTTGTGATGATG 540
Db      481 AATATGACAAATGAACCTGTGCTGATGCCATGCGGCGAAAGGTTTGTGATGATG 540
QY      541 TGCAGAGAGTTGAGGCTTACATGGAATCCATGCTTCACTCCCGGAACTATGAGC 600
Db      541 TGCAGAGAGTTGAGGCTTACATGGAATCCATGCTTCACTCCCGGAACTATGAGC 600
QY      601 CGGATTTCAATGCTGCCAACCAAGAGAGCGTCCCAACAGCTCATCAAGGGACAAAGC 660
Db      601 CGGATTTCAATGCTGCCAACCAAGAGAGCGTCCCAACAGCTCATCAAGGGACAAAGC 660
QY      661 AAGAGCAAGTTCAACCAATCATCAAGAGATCAAGGCGTTTAAAGGACCAACAAAGTGG 720
Db      661 AAGAGCAAGTTCAACCAATCATCAAGAGATCAAGGCGTTTAAAGGACCAACAAAGTGG 720
QY      721 ACAAGGTGTTGACTGTGGAATGCGCAACAGAGAGTACAGTAAATTTGTTGGGCC 780
Db      721 ACAAGGTGTTGACTGTGGAATGCGCAACAGAGAGTACAGTAAATTTGTTGGGCC 780
QY      781 TTAATGACACCAATGAGAAATCTTTGGCTGCTGTGACAGAAATGAAGCTGAGATTTCTC 840
Db      781 TTAATGACACCAATGAGAAATCTTTGGCTGCTGTGACAGAAATGAAGCTGAGATTTCTC 840
QY      841 CTTCACCTTTGATGTCATTTGCTTGTGTTATGAAAAATGTTCTTTCAATTAAGAGGCC 900
Db      841 CTTCACCTTTGATGTCATTTGCTTGTGTTATGAAAAATGTTCTTTCAATTAAGAGGCC 900
QY      901 CTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTCCATCGGAGAGAACCTTTGATTG 960
Db      901 CTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTCCATCGGAGAGAACCTTTGATTG 960
QY      961 GTGAGATGACTTCAAGAGTGTCAAGCCAAATGAATCTGTGTGTGTTGATTTCTTG 1020
Db      961 GTGAGATGACTTCAAGAGTGTCAAGCCAAATGAATCTGTGTGTGTTGATTTCTTG 1020
QY      1021 TGGGGGCTGGTATCAAGCCAAATCTATATGTCAGTTACAAACATCTGGGAAACATGATG 1080
Db      1021 TGGGGGCTGGTATCAAGCCAAATCTATATGTCAGTTACAAACATCTGGGAAACATGATG 1080

```

```

QY      1081 GTATGAATCTTTGCGCTCCACAACTTCCGTTCCAGAAATCTCCAAAGCAACGTTG 1140
Db      1081 GTATGAATCTTTGCGCTCCACAACTTCCGTTCCAGAAATCTCCAAAGCAACGTTG 1140
QY      1141 TTGATGATATGTTCAACAGCAATGCGATCTCTATGAGCTGTGTGAACATCCAGCCATG 1200
Db      1141 TTGATGATATGTTCAACAGCAATGCGATCTCTATGAGCTGTGTGAACATCCAGCCATG 1200
QY      1201 TTGTTGTTATTAAGATATGCTTACGTATGAGGAGACAGCAAGAGACCATGATGATGACA 1260
Db      1201 TTGTTGTTATTAAGATATGCTTACGTATGAGGAGACAGCAAGAGACCATGATGATGACA 1260
QY      1261 CTTGAGATATATTCATGAGGTGAGAAAGACACCATTTGTTTGACAAACATGCGAGATT 1320
Db      1261 CTTGAGATATATTCATGAGGTGAGAAAGACACCATTTGTTTGACAAACATGCGAGATT 1320
QY      1321 CCTCTTATGCTGCTCTTATTTATCTTGACCTTGTGCTTCTTGTGAGCTCAGACTATAGA 1380
Db      1321 CCTCTTATGCTGCTCTTATTTATCTTGACCTTGTGCTTCTTGTGAGCTCAGACTATAGA 1380
QY      1381 TCGAGTTTAAAGCTGAAATAGGAGAAATTCACATTCACCCAGTGTGTAATGCAATCC 1440
Db      1381 TCGAGTTTAAAGCTGAAATAGGAGAAATTCACATTCACCCAGTGTGTAATGCAATCC 1440
QY      1441 TCAGCTACCTCAACCAAGGCTCTGTGTTCACCGGATACACAGTGTGAATGCAATGT 1500
Db      1441 TCAGCTACCTCAACCAAGGCTCTGTGTTCACCGGATACACAGTGTGAATGCAATGT 1500
QY      1501 CAAAGCAGCTGCAATGCTGTGAAAAATATATGAGAGGCTTGTGATTTGGCCCAAGAGA 1560
Db      1501 CAAAGCAGCTGCAATGCTGTGAAAAATATATGAGAGGCTTGTGATTTGGCCCAAGAGA 1560
QY      1561 AATAATGATTTCTGAGTACAAAGTGAAGATGGGACCGAAGAAATATATGTTGGGGTATG 1620
Db      1561 AATAATGATTTCTGAGTACAAAGTGAAGATGGGACCGAAGAAATATATGTTGGGGTATG 1620
QY      1621 CCTAGCTGAATGTTTATGTTATATATATATGTTTGTCTTATTAATTTTGAAGTGAATGA 1680
Db      1621 CCTAGCTGAATGTTTATGTTATATATATATATGTTTGTCTTATTAATTTTGAAGTGAATGA 1680
QY      1681 TGCATCAGCTTCAATTAATGCTTTAGAGCGGGGCAATTTCTGTTACTAGAAACATGAATG 1740
Db      1681 TGCATCAGCTTCAATTAATGCTTTAGAGCGGGGCAATTTCTGTTACTAGAAACATGAATG 1740
QY      1741 AATGTAGTAAATTTTGTGT 1760
Db      1741 AATGTAGTAAATTTTGTGT 1760

```

RESULT 2
 AY038802 1739 bp mRNA linear PLN 16-JUN-2001
 LOCUS
 DEINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 1739)
 AUTHORS
 Hitz,W.D., Carlson,T.J., Kerr,P. and Sebastian,S.
 Biochemical and Molecular Characterization of a Mutation that
 Confers a Decreased Raffinose and Phytic Acid Phenotype on
 Soybean Seeds
 Unpublished
 2 (bases 1 to 1739)
 REFERENCE
 Carlson,T.J. and Hitz,W.D.
 Direct Submission
 Submitted (06-JUN-2001) Crop Genetics, DuPont Co., P.O.Box 80402,
 Wilmington, DE 19880-0402, USA

FEATURES

source

Location/Qualifiers
 1. 1739
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Wye"
 /db_xref="taxon:3847"
 /dev_stage="developing embryo"
 11. 1543
 /EC_number="5.5.1.4"
 /function="cyclizes glucose 6-phosphate to
 L-myo-inositol-1-phosphate"
 /note="mIPS-1A"
 /codon_start=1
 /product="myo-inositol-1-phosphate synthase"
 /protein_id="AAK72098.1"
 /db_xref="GI:1476446"

/translation="MPIENPKYECPNVKTETELIQSVNYETTELVENHNGTYOMIV
 KKSXYEERKTHIHYPLKGLVLMVGMNGSLTGVIANREGISWATOKIQAANYF
 GSLTOSAIRVGSFOGEEIYAPFKSLIPVNPDDIVFGMDISNMNLADAMRAKVED
 IDLOKOLRYMESMLPLBGIYDPDPIAANOERANVICTROEOVOQIHKIKAFKE
 ATKVDKVVLTANTERYSNLVVGLNDTMENTLAVDREAEISPTLYALACWENV
 PFINSQPTFVPGILDLAIARNTLIGDDPFGSQTKMSVLDFLVGAGIKPTSLVS
 YNHLGNNDGNLISAPQTFRSKEISKSNVDDMVNSNALIYERPEHDPHYVVKYFV
 GSKRAMDEYTSBIFMGKSTIYLAHNTCEDSLAAPIIDLVLAEISLRIEKENE
 GKPSHFHPVATILSYLTKARIVPDPPTPVNVALSKORAMLENTMRACVGLAPENMILE
 YK"

ORIGIN

Query Match 97.6%; Score 1717; DB 8; Length 1739;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GTGAAAAATATGTTTCATCGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAATACCA 103
 DB 1 GTGAAAAATATGTTTCATCGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAATACCA 60
 QY 104 GACTGAGATTCAGTCCGTGTACAACTACGAAACCAACCACTGTTACAGAGAACAGGAA 163
 DB 61 GACTGAGATTCAGTCCGTGTACAACTACGAAACCAACCACTGTTACAGAGAACAGGAA 120
 QY 164 TGGCACTTACATGATGATTCACAAACCAATCTGTCAATAAGAAATTTAAACCAACAT 223
 DB 121 TGGCACTTACATGATGATTCACAAACCAATCTGTCAATAAGAAATTTAAACCAACAT 180
 QY 224 CCATGTTCTTAATTAAGGGGTAAATGCTTGTGGGTGGGAGAACACGGCTCAACCT 283
 DB 181 CCATGTTCTTAATTAAGGGGTAAATGCTTGTGGGTGGGAGAACACGGCTCAACCT 240
 QY 284 CACCGGTGTGTTATGCTTAAACGAGAGGCAATTTGATGGGCTACAAAGAACAAATTTCA 343
 DB 241 CACCGGTGTGTTATGCTTAAACGAGAGGCAATTTGATGGGCTACAAAGAACAAATTTCA 300
 QY 344 ACAAGCAATTAATTTGGCTCCCTCAACCAAGCTCAGCTATCCAGATTGGCTCTTCA 403
 DB 301 ACAAGCAATTAATTTGGCTCCCTCAACCAAGCTCAGCTATCCAGATTGGCTCTTCA 360
 QY 404 GAGAGAGAAATCTATGCCCATTTCAAGAGCTGCTTCAATGTTAACCTTGACGACAT 463
 DB 361 GAGAGAGAAATCTATGCCCATTTCAAGAGCTGCTTCAATGTTAACCTTGACGACAT 420
 QY 464 TGTGTTTGGGGATGAGGATATCAGCAACATGAACTGTGCTGAATCCATGGCCAGGCAAA 523
 DB 421 TGTGTTTGGGGATGAGGATATCAGCAACATGAACTGTGCTGAATCCATGGCCAGGCAAA 480
 QY 524 GGTGTTTGAATGATTTTGCAGAGAGAGTTGAGGCTTACATGGAATCCATGCTTCACT 583
 DB 481 GGTGTTTGAATGATTTTGCAGAGAGAGTTGAGGCTTACATGGAATCCATGCTTCACT 540
 QY 584 CCCCGGAATCTATGACCCGAGATTTCAATGTGCTCAACCAAGAGAGCGTGGCCAAACGT 643
 DB 541 CCCCGGAATCTATGACCCGAGATTTCAATGTGCTCAACCAAGAGAGCGTGGCCAAACGT 600
 QY 644 CATCAAGGCAACCAAGAGAGAGTTCACAAATCATCAAGAGATCAAGGCGTTTAA 703

DB 601 CATCAAGGCAACCAAGAGAGAGTTCACAAATCATCAAGAGATCAAGGCGTTTAA 660
 QY 704 GGAAGCCACCAAGATGAGCAAGGTGTTACTGTGAGCTGCCAACAGAGAGTACAG 763
 DB 661 GGAAGCCACCAAGATGAGCAAGGTGTTACTGTGAGCTGCCAACAGAGAGTACAG 720
 QY 764 TAATTTGTTGGGGCTTAATGACCAATGAGAAATCTTGGCGCTGTGAGCAAGAA 823
 DB 721 TAATTTGTTGGGGCTTAATGACCAATGAGAAATCTTGGCGCTGTGAGCAAGAA 780
 QY 824 TAGAGCTGAGATTTCTCTTCCACCTTGATGCCATTCCTTGTTATGAGAAATGTTCC 883
 DB 781 TAGAGCTGAGATTTCTCTTCCACCTTGATGCCATTCCTTGTTATGAGAAATGTTCC 840
 QY 884 TTTCAATTAATGAGAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGTCATGCC 943
 DB 841 TTTCAATTAATGAGAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGTCATGCC 900
 QY 944 GAGGAACACTTTGATTTGTGTGAGATGACTTCAAGAGTGTGACCAAAATGAAATCTGT 1003
 DB 901 GAGGAACACTTTGATTTGTGTGAGATGACTTCAAGAGTGTGACCAAAATGAAATCTGT 960
 QY 1004 GTTGGTTGATTTCTTGTGGGGCTGTGTATCAAGCCACATCTATAGTCAATTAACCA 1063
 DB 961 GTTGGTTGATTTCTTGTGGGGCTGTGTATCAAGCCACATCTATAGTCAATTAACCA 1020
 QY 1064 TCTGGGAAACCAATGATGATGATTAATCTTCCGCTCCACAACTTCCGTTCCAGAGAAAT 1123
 DB 1021 TCTGGGAAACCAATGATGATGATTAATCTTCCGCTCCACAACTTCCGTTCCAGAGAAAT 1080
 QY 1124 CTCGAAGAGCAACGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1183
 DB 1081 CTCGAAGAGCAACGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 QY 1184 TGAACATCCAGACCAATGTTGTTGTTATTAAGATGATGATGATGATGATGATGATGATG 1243
 DB 1141 TGAACATCCAGACCAATGTTGTTGTTATTAAGATGATGATGATGATGATGATGATGATG 1200
 QY 1244 AGCCATGATGATGATCACTTCCAGAGATTAATCAATGGGTGAGAAAGAGCAACATTTTGTGA 1303
 DB 1201 AGCCATGATGATGATCACTTCCAGAGATTAATCAATGGGTGAGAAAGAGCAACATTTTGTGA 1260
 QY 1304 CAACACATGCGAGATTCCTCTTACGCTGCTCTTATATCTTGAATGATGATGATGATGATG 1363
 DB 1261 CAACACATGCGAGATTCCTCTTACGCTGCTCTTATATCTTGAATGATGATGATGATGATG 1320
 QY 1364 TGAGCTCAGCATTAGATGATGATGATTAAGCTGAATGAGGAGAAATTCACATCAATTTCA 1423
 DB 1321 TGAGCTCAGCATTAGATGATGATGATTAAGCTGAATGAGGAGAAATTCACATCAATTTCA 1380
 QY 1424 CCCAGTTGTACATCTTACGCTTACCTCAACAGGCTCCTCTGTGTTCCACCGGGATACAC 1483
 DB 1381 CCCAGTTGTACATCTTACGCTTACCTCAACAGGCTCCTCTGTGTTCCACCGGGATACAC 1440
 QY 1484 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1543
 DB 1441 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
 QY 1544 TGGATTTGGCCCAAGAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATG 1603
 DB 1501 TGGATTTGGCCCAAGAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATG 1560
 QY 1604 TAATATAGTTGGGGATAGCTTACGATGATGATGATGATGATGATGATGATGATGATGAT 1663
 DB 1561 TAATATAGTTGGGGATAGCTTACGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 QY 1664 TTGCAAGTGAATTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1723
 DB 1621 TTGCAAGTGAATTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 QY 1724 TACTAGGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1760

DB 1681 TACTAGGACATGAATGAATGATATAATTTTGTGT 1717

RESULT 3
BD075269

LOCUS
DEFINITION Soybean plant producing seeds with reduced levels of raffinose saccharides and phytic acid.

ACCESSION BD075269

VERSION BD075269.1 GI:22620872

KEYWORDS JP 2001519665-A/4.

SOURCE unidentified

ORGANISM unidentified

REFERENCE 1 (bases 1 to 1533)
Hitz, W.D. and Sebastian, S.A.
Soybean plant producing seeds with reduced levels of raffinose saccharides and phytic acid
Patent: JP 2001519665-A 4 23-OCT-2001;

JOURNAL
EI DU PONT DE NEMOURS AND CO

COMMENT
OS Soybean line LR33
PN JP 2001519665-A/4
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN
PC C12N15/52, C12N15/82, C12N15/11, C12N9/90, A01H5/00 CC
Strandness: Double;
CC Topology: Linear;
CC Soybean plant producing seeds with reduced levels of raffinose
CC saccharides
FH key and phytic acid
FT CDS 1:1533.

FEATURES
source location/Qualifiers
1..1533
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 84.2%; Score 1482; DB 6; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 54 ATGTTATGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCCGAGACTGAGATT 113
DB 1 ATGTTATGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCCGAGACTGAGATT 60

QY 114 CAGTCGCTGAACAACCTACGGAACCTGTTACGAGAAAGAGAAATGGACCTAT 173
DB 61 CAGTCGCTGAACAACCTACGGAACCTGTTACGAGAAAGAGAAATGGACCTAT 120

QY 174 CAGTCGCTGAACAACCTACGGAACCTGTTACGAGAAAGAGAAATGGACCTAT 233
DB 121 CAGTCGCTGAACAACCTACGGAACCTGTTACGAGAAAGAGAAATGGACCTAT 180

QY 234 AAATTTGGGGTAATGCTTTGGGGTGGGGTGGGAAACAAGGCTCAACCTCACGGTGGT 293
DB 181 AAATTTGGGGTAATGCTTTGGGGTGGGGTGGGAAACAAGGCTCAACCTCACGGTGGT 240

QY 294 GTTATTGCTAACGAGAGGGCATTTGATGGGCTAACAAAGGACAAAGTTCACAAAGCAAT 353
DB 241 GTTATTGCTAACGAGAGGGCATTTGATGGGCTAACAAAGGACAAAGTTCACAAAGCAAT 300

QY 354 TACTTTGGCTCCCTCAACCCCAAGCTCAGCTATCCGAGTTGGGTCTTTCCAGGAGAGAA 413
DB 301 TACTTTGGCTCCCTCAACCCCAAGCTCAGCTATCCGAGTTGGGTCTTTCCAGGAGAGAA 360

QY 414 ATCTATGCCCCCATTTCAAGAGCTGCTTTCCAAATGTTTAACTTGACGACATTTGTTGGG 473
DB 361 ATCTATGCCCCCATTTCAAGAGCTGCTTTCCAAATGTTTAACTTGACGACATTTGTTGGG 420

QY 474 GGATGGATATACGACATGAACTGTGCTGATGTCATGGCCAGGGCAAGGTGTTTAC 533
DB 421 GGATGGATATACGACATGAACTGTGCTGATGTCATGGCCAGGGCAAGGTGTTTAC 480

QY 534 ATCGATTTTGGCAGAGCACTTGAAGCTTTTACATGAAATCAATGCTTCCACTCCCGGAATC 593
DB 481 ATCGATTTTGGCAGAGCACTTGAAGCTTTTACATGAAATCAATGCTTCCACTCCCGGAATC 540

QY 594 TATGACCCGGATTTTATGCTGCTGCCAACCAAGAGAGCGTGCACAAAGTTCATCAAGGAC 653
DB 541 TATGACCCGGATTTTATGCTGCTGCCAACCAAGAGAGCGTGCACAAAGTTCATCAAGGAC 600

QY 654 ACAAGCAAGAGCAAGTTCAACCAATCATCAAGACATCAAGGCTTTAAGAACCAAC 713
DB 601 ACAAGCAAGAGCAAGTTCAACCAATCATCAAGACATCAAGGCTTTAAGAACCAAC 660

QY 714 AAAGTGAACAAGTGTGTGACTGTGACCTGCCAACACAGAGAGGTACAGTAATTTGGTT 773
DB 661 AAAGTGAACAAGTGTGTGACTGTGACCTGCCAACACAGAGAGGTACAGTAATTTGGTT 720

QY 774 GTGGGCTTTAATGACACCATGAGAAATCTTGGCTGTGAGACAGAAATGAGGCTGAG 833
DB 721 GTGGGCTTTAATGACACCATGAGAAATCTTGGCTGTGAGACAGAAATGAGGCTGAG 780

QY 834 ATTTCTCCTTCCACCTTGTATGCCATTTGCTGTGTATGAAATGTTCTTTTCAATAT 893
DB 781 ATTTCTCCTTCCACCTTGTATGCCATTTGCTGTGTATGAAATGTTCTTTTCAATAT 840

QY 894 GGAAGCCTCAGAACACTTTTGTATCCAGAGGCTGATTTGATCTTGGCATGCGGAGAACCT 953
DB 841 GGAAGCCTCAGAACACTTTTGTATCCAGAGGCTGATTTGATCTTGGCATGCGGAGAACCT 900

QY 954 TTGATTTGGAGATGACCTTCAAGAGTGTGTCAGACCAAAATGAATCTGTGTGTTGAT 1013
DB 901 TTGATTTGGAGATGACCTTCAAGAGTGTGTCAGACCAAAATGAATCTGTGTGTTGAT 960

QY 1014 TTCTTTGGGGGCTGTATCAAGGCCAACATCTATAGTCAGTTACAAACATCTGGGAAAC 1073
DB 961 TTCTTTGGGGGCTGTATCAAGGCCAACATCTATAGTCAGTTACAAACATCTGGGAAAC 1020

QY 1074 AATGATGATGATATCTTTGGCTCCACAACTTCCGTTCCAGAAATCTCCAAAGAC 1133
DB 1021 AATGATGATGATATCTTTGGCTCCACAACTTCCGTTCCAGAAATCTCCAAAGAC 1080

QY 1134 AAGCTTTGATGATATGATGTCACAGCAATGCCATCTCTATGAGCCGTGTAATCCA 1193
DB 1081 AAGCTTTGATGATATGATGTCACAGCAATGCCATCTCTATGAGCCGTGTAATCCA 1140

QY 1194 GACCATGTTGTTGTTAATGATGTCCTTACGTAGGGGACAGCAAGAGCAATGAT 1253
DB 1141 GACCATGTTGTTGTTAATGATGTCCTTACGTAGGGGACAGCAATGAGCAATGAT 1200

QY 1254 GAGTACACTTCAAGATATTCATGGGTGAAAGAGACACATTTGTTTGCACAAACATGC 1313
DB 1201 GAGTACACTTCAAGATATTCATGGGTGAAAGAGACACATTTGTTTGCACAAACATGC 1260

QY 1314 GAGGATCCCTCTTACGCTCTCTATTAATCTTGAACCTTGGTCTTGTGAGCTCAGC 1373
DB 1261 GAGGATCCCTCTTACGCTCTCTATTAATCTTGAACCTTGGTCTTGTGAGCTCAGC 1320

QY 1374 ACTAGAAATCGAATTTAAAGCTGAAAGAGGAAATTTCACTCAATCCACCAAGTTGT 1433
DB 1321 ACTAGAAATCGAATTTAAAGCTGAAAGAGGAAATTTCACTCAATCCACCAAGTTGT 1380

QY 1434 ACCATCTCAGCTTACCTCAACAGGCTCTCTGTTTCCACCGGGTACACCAAGTGTGAT 1493
DB 1381 ACCATCTCAGCTTACCTCAACAGGCTCTCTGTTTCCACCGGGTACACCAAGTGTGAT 1440

QY 1494 GCATTTGTCAAAGCAGCTGCAATGCTGGAATAATTAAGAGGCTTGTGTTGATTTGGCC 1553
DB 1441 GCATTTGTCAAAGCAGCTGCAATGCTGGAATAATTAAGAGGCTTGTGTTGATTTGGCC 1500

QY	1554	CCAGAGATAACATGATCTTCGAGACAAAGCA	1586
Db	1501	CCAGAGATTAACATGATCTTCGAGTACAAAGTGA	1533
RESULT 4			
LOCUS	AF293970	1791 bp	mRNA
DEFINITION	AF293970 Glycine max myo-inositol-3-phosphate synthase (MIPS) mRNA, complete cds.		
ACCESSION	AF293970		
VERSION	AF293970.1	GI:13936690	
KEYWORDS	Glycine max (soybean)		
SOURCE	Glycine max (soybean)		
ORGANISM	Glycine max (soybean)		
REFERENCE	Glycine max (soybean)		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.		
TITLE	1 (bases 1 to 1791)		
JOURNAL	Hegeman, C.E., Good, L.L. and Grabau, E.A.		
MELINE	Expression of D-myo-inositol-3-phosphate synthase in soybean.		
PUBMED	Implications for phytic acid biosynthesis		
REFERENCE	Plant Physiol. 125 (4), 1941-1948 (2001)		
AUTHORS	2 (bases 1 to 1791)		
TITLE	Hegeman, C.E., Good, L.L. and Grabau, E.A.		
JOURNAL	Direct Submission		
MELINE	Submitted (07-AUG-2000) Plant Pathology, Physiology, and Weed Science, Virginia Tech, Fralin Biotechnology Center, Blacksburg, VA 24060, USA		
FEATURES	Location/Qualifiers		
source	1..1791		
	/organism="Glycine max"		
	/mol_type="mRNA"		
	/cultivar="Williams 82"		
	/db_xref="taxon:3847"		
	/dev_stage="immature cotyledons"		
	1..1791		
	/gene="MIPS"		
	62..1594		
	/gene="MIPS"		
	/note="isomerase"		
	/codon_start=1		
	/product="myo-inositol-3-phosphate synthase"		
	/protein_id="AAK4986.1"		
	/db_xref="GI:13936691"		
	/translation="MFLENFVCECPNVKYTEIQTQSYVNNETTELVEHNRNGTYQWIVKPKSKYAEFKTNIHVPKGLVNLVGMGNGNSLTGVIYANNEGISMATKXIQQANFPGSLQASAIKRVGSEIYAPFKSLIPWVDDIVFGMDISNNLADMAKAPFDIDLOKQMPVWESMLPLPGIYDPPFIANOBERANNVYKTKGBOVOOIIDIKAFKREATKDKQVVLMTANTERSNLTVGINDMENTLAAVDRNEAETSPSTLYAIACMENVVPFNGSPQNTVPEGLIDAIKRNLTIGSDPKSKQTKKSLTVDPVUGAGIKPSTIYSYHGLKNDGKMLSAQVTKRSKISKSNVDVNSNALVPGEPHVVIVKVPVYVYNSGRANDVETYSIEFMGKSTVLANCTEDSLAAPITLIDVLAELSTRLEFPAENBEGKFSFHEVATILSYLTAKPLVPPGTPEVNALSKORAMLENIMRACVGLAPENNMLEYK"		
ORIGIN			
Query Match	82.6%	Score 1454;	DB 8; Length 1791;
Best Local Similarity	99.7%	Pred. No. 0;	
Matches 1754;	Conservative 0;	Mismatches 6;	Indels 0; Gaps 0;
QY	1	CTCTCTTTATTCCTTTGTTATTCATTCTTATCTTTGGAATAATATGTTCA	60
Db	9	CTCTCTTTATTCCTTTGTTATTCATTCTTATCTTTGGAATAATATGTTCA	68
QY	61	TCGAGAAATTTAAAGGTGAGTGTCTTAATGTGAAGTACACGAGACTGAGATTAGTCGC	120
Db	69	TCGAGAAATTTAAAGGTGAGTGTCTTAATGTGAAGTACACGAGACTGAGATTAGTCGC	128
QY	121	TGTACAACCTACGAAACACCGAGACTTGTTCACGAGACAGAAATGGACCTATACGTGGA	180

Db	129	TGTCAACTACGAAACCAACCGA	CTTGTTCACGAGAACGAAATGGCACTTACAGTGA	188
Qy	181	TTGTCAAAACCAAACTGTGTCA	AATACGAATTTAAACCAACATCCATGTTCTTAAATTAG	240
Db	189	TTGTCAAAACCAAACTGTGTCA	AATACGAATTTAAACCAACATCCATGTTCTTAAATTAG	248
Qy	241	GGGTATGCTTGTGGGTGGGGT	GGAACCAACGCTCAACCTTCAACCGTGTGTATTG	300
Db	249	GGGTATGCTTGTGGGTGGGGT	GGAACCAACGCTCAACCTTCAACCGTGTGTATTG	308
Qy	301	CTAACCGAGAGGCAATTCATG	GGGCTACAAAGAGCAAGATTCAACCAACTTACTTGG	360
Db	309	CTAACCGAGAGGCAATTCATG	GGGCTACAAAGAGCAAGATTCAACCAACTTACTTGG	368
Qy	361	GCTCCCTACCCAAAGCCTCAG	CTATCCGAGTTGGGTCCCTTCACAGGAGGAAATCTATG	420
Db	369	GCTCCCTACCCAAAGCCTCAG	CTATCCGAGTTGGGTCCCTTCACAGGAGGAAATCTATG	428
Qy	421	CCCCATTCAAGAGCTGCTTCA	ATGTTAAACCTTGACACATTTGTGTTGGGGATGGG	480
Db	429	CCCCATTCAAGAGCTGCTTCA	ATGTTAAACCTTGACACATTTGTGTTGGGGATGGG	488
Qy	481	ATATAGCAATGAACCTGGCTG	ATATGCAAGGCAAGGCAAGGATGTTTACATCGATT	540
Db	489	ATATAGCAATGAACCTGGCTG	ATATGCAAGGCAAGGCAAGGATGTTTACATCGATT	548
Qy	541	TGCAGAACATTTGAGGCTTAC	ATGAGATCATGCTTCCATCCCGGAATCTATGACC	600
Db	549	TGCAGAACATTTGAGGCTTAC	ATGAGATCATGCTTCCATCCCGGAATCTATGACC	608
Qy	601	CGGATTCATTTGCTGCCAAC	CAAGAGAGCGTGCACACATCATCAAGGCAACAGC	660
Db	609	CGGATTCATTTGCTGCCAAC	CAAGAGAGCGTGCACACATCATCAAGGCAACAGC	668
Qy	661	AAGAGCAAGTTCACAAATTC	ATCAAAAGCGTTTAAAGAAAGCCCAAAAGTG	720
Db	669	AAGAGCAAGTTCACAAATTC	ATCAAAAGCGTTTAAAGAAAGCCCAAAAGTG	728
Qy	721	ACAAGGTGTTGTACTGTGGA	CTGCCAACAGAGAGTACAGTAAATTTGGTTGGGCC	780
Db	729	ACAAGGTGTTGTACTGTGGA	CTGCCAACAGAGAGTACAGTAAATTTGGTTGGGCC	788
Qy	781	TTAATGACACATGAGAGATCT	CTTGGCTGTGTGACAGAAATGAGGCTGAGATTCTC	840
Db	789	TTAATGACACATGAGAGATCT	CTTGGCTGTGTGACAGAAATGAGGCTGAGATTCTC	848
Qy	841	CTTCCACTTGTATGCCATTG	CTTGTGTATGAAAAATGTTCCCTTCAATTAATGAAGCC	900
Db	849	CTTCCACTTGTATGCCATTG	CTTGTGTATGAAAAATGTTCCCTTCAATTAATGAAGCC	908
Qy	901	CTCAGAACCTTTTGTATCAG	GGGCTGATTAATCTTGGCCATGCGAGAGAACTTTGATTG	960
Db	909	CTCAGAACCTTTTGTATCAG	GGGCTGATTAATCTTGGCCATGCGAGAGAACTTTGATTG	968
Qy	961	GTCAGATGACTTCAAGATG	GTGACCAAAATGAATCTGTGTGGTTGATTTCTTG	1020
Db	969	GTCAGATGACTTCAAGATG	GTGACCAAAATGAATCTGTGTGGTTGATTTCTTG	1028
Qy	1021	TGGGGGCTGTATCAAGCCAC	ATCTATATGTCAGTTACAAACCATCTGGGAAACAAATGATG	1080
Db	1029	TGGGGGCTGTATCAAGCCAC	ATCTATATGTCAGTTACAAACCATCTGGGAAACAAATGATG	1088
Qy	1081	GATGAATCTTTGGGCTCCACA	AACTTTCCGTTCCAGAGAAATCTCCAAAGCAACGTTG	1140
Db	1089	GATGAATCTTTGGGCTCCACA	AACTTTCCGTTCCAGAGAAATCTCCAAAGCAACGTTG	1148
Qy	1141	TTGATGATATGTCACACAG	CAATCCATCTATAGAGCTGTGAAACATTCACAGCATG	1200
Db	1149	TTGATGATATGTCACACAG	CAATCCATCTATAGAGCTGTGAAACATTCACAGCATG	1208
Qy	1201	TTGTTGTTATTAAGTATGCTT	ATACGTAGAGGACAGCAAGAGCCATGATGATGATCA	1266

Db	1209	TTGTTGTTAATTAAGTATGTGCTTACGTAAGGAGGACAGCAAGAGGCGCATGGATGAGTACA	1268
QY	1261	CTTCAGAGATATTTATGSGGTGGAAAAGACACCATTTGTTTCACACACATGCGAGATT	1320
Db	1269	CTTCAGAGATATTTATGSGGTGGAAAAGACACCATTTGTTTCACACACATGCGAGATT	1328
QY	1321	CCCTCTTAGCGTGCCTATTAATCTTGGACCTTGGTCCCTCTTGCTGAGCTCAGACTAGAA	1380
Db	1329	CCCTCTTAGCGTGCCTATTAATCTTGGACCTTGGTCCCTCTTGCTGAGCTCAGACTAGAA	1388
QY	1381	TCGAGTTTAAAGCTGAAAATGAGGGAATAATTCACATCTCCACCCAGTTGCTACATCC	1440
Db	1389	TCGAGTTTAAAGCTGAAAATGAGGGAATAATTCACATCTCCACCCAGTTGCTACATAC	1448
QY	1441	TCAGCTTACCTCAACAAAGGCTCTCTGTGTTCCACCGGGTCAACCAAGTGTGAATGCATTGT	1500
Db	1449	TCAGCTTACCTCAACAAAGGCTCTCTGTGTTCCACCGGGTCAACCGGGTGTGAATGCATTGT	1508
QY	1501	CAAGCGACCGCGCAATGCTGSAANAACATAAGAGGGCTTGATTTGGCCCCACAGA	1560
Db	1509	CAAGCGACCGCGCAATGCTGSAANAACATAAGAGGGCTTGATTTGGCCCCACAGA	1568
QY	1561	ATAACATGATTTCTGAGATCAAGTGAAGCATGGGACCGCAAAATATATATTAATTTGGGGTAG	1620
Db	1569	ATAACATGATTTCTGAGATCAAGTGAAGCATGGGACCGCAAAATATATATTAATTTGGGGTAG	1628
QY	1621	CTTAGCTGAATGTTTATATGTTAATATATATGTTTCTTATTAATTTTGCACAGTATTTGAA	1680
Db	1629	CTTAGCTGAATGTTTATATGTTAATATATATGTTTCTTATTAATTTTGCACAGTATTTGAA	1688
QY	1681	TGCAATCAGCTTCATTAATGCTTTTAGCGGGGGCATATTCGTGTTTACTAGAAACATGAATG	1740
Db	1689	TGCAATCAGCTTCATTAATGCTTTTAGCGGGGGCATATTCGTGTTTACTAGAAACATGAATG	1748
QY	1741	AATGATGATTAATTTTGCT 1760	
Db	1749	AATGATGATTAATTTTGCT 1768	

RESULT 5	
AP006418	
LOCUS	62713 bp DNA linear PLN 22-JUL-2003
DEFINITION	Lotus corniculatus var. japonicus genomic DNA, chromosome 6, clone:lj1t23a02, TM0306, complete sequence.
ACCESSION	AP006418
VERSION	
KEYWORDS	AP006418.1 GI:31581049
ORGANISM	HTG.
ORGANISM	Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM	Lotus corniculatus var. japonicus
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
REFERENCE	1
AUTHORS	Asamizu,E., Kato,T., Sato,S., Nakamura,Y., Kaneko,T. and Tabata,S.
TITLE	Structural Analysis of a Lotus japonicus Genome. IV. Sequence Features and Mapping of seventy-three TAC clones which cover the 7.5 Mb Regions of the Genome
JOURNAL	DNA Res. (2003) In press
REFERENCE	2 (bases 1 to 62713)
AUTHORS	Sato,S.
TITLE	Direct Submission
JOURNAL	Submitted (07-MAY-2003) Shussei Sato, Karusa DNA Research Institute, Department of Plant Gene Research; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail:issato@kazusa.or.jp, URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935 (ex.2337), Fax:81-438-52-3934)
FEATURES	location/Qualifiers
SOURCE	1..62713

	/chromosome=“6” /clone=LJTJ23A02” /clone_lib=“LJT_library” /note=“TAC clone:TW0306-synonym: Lotus japonicus”
ORIGIN	
Query Match Best Local Similarity 100.0%; Pred. No. 1.2e-24; Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
DQ	1140 GTTGATGATNTGGTCACAAGCAATGCCATCCTTTATGAAGCGGTGTAAACATCCAGAACCATT 1199
Dd	30144 GTTGATGATNTGGTCACAAGCAATGCCATCCTCTTATGAAGCGGTGTAAACATCCAGAACCATT 30203
OY	1200 GTTGT 1204
Dd	30204 GTTGT 30208

RESULT 6	2994 bp	DNA	linear	PLN 02-JUL-2001
AF282263				
LOCUS				
DEFINITION	Phaseolus vulgaris 1L-myo-inositol-1-phosphate synthase gene,			
ACCESSION	partial cds.			
VERSION	AF282263			
KEYWORDS	AF282263.1	GI:14582466		
SOURCE				
ORGANISM	Phaseolus vulgaris			
	Phaseolus vulgaris			
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	rosids; eurosids I, Fabales; Fabaceae; Papilionoideae; Phaseoleae;			
	Phaseolus.			
REFERENCE	1 (bases 1 to 2994)			
AUTHORS	Johnson,M.D., Lackey,K.H. and Pope,P.M.			
TITLE	Subcellular localization of 1L-myo-inositol-1-phosphate synthase			
JOURNAL	unpublished			
REFERENCE	2 (bases 1 to 2994)			
AUTHORS	Pope,P.M. and Johnson,M.D.			
TITLE	Direct Submissions			
JOURNAL	Submitted (24-JUN-2000) Biological Sciences, The University of			
	Alabama, P.O. Box 870344, Tuscaloosa, AL 35487, USA			
FEATURES	Location/Qualifiers			
SOURCE	1..2994			

3.4%; Score 59; DB 8; Length 2994;
 Best Local Similarity 100.0%; Pred. No. 3 8e-21;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1134 AACGTTGATGATGATGTCACAGCAATGCATCTCTTATGAGCTGGTGAACATCC 1192

Db	2339	AACGTTGTGATGATATGTCACAACGAATGCCATCTCTAATAGACCTGGTGAATCC	2397
RESULT 7			
AY005128		1391 bp mRNA linear	PLN 30-JAN-2001
LOCUS			
DEFINITION	Actinidia arguta myo-inositol-1-phosphate synthase mRNA, complete cds.		
ACCESSION	AY005128		
VERSION	AY005128.1 GI:12597482		
SOURCE			
ORGANISM	Actinidia arguta Actinidia arguta Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; asterids; Ericales; Actinidiaceae; Actinidia. 1 (bases 1 to 1391) Klaesges,K., Fitzgerald,A. and Moodie,M. Synthesis of inositol in kiwifruit Unpublished 2 (bases 1 to 1391) Klaesges,K., Fitzgerald,A. and Moodie,M. Direct Submission Submitted (18-JUL-2000) Hortresearch, 120 Mt Albert Rd., Auckland, New Zealand		
FEATURES			
Source	Location/Qualifiers		
	1..1391		
	/organism="Actinidia arguta"		
	/mol_type="mRNA"		
	/db_xref="taxon:64478"		
	/cissue.type="young mature leaf"		
CDS	1..1095		
	/codon_start=1		
	/product="myo-inositol-1-phosphate synthase"		
	/protein_id="AAF97409.1"		
	/db_xref="GI:12597483"		
	/translation="MNLADMARARVPVIDQKLRPYMESMPPEGLYEPDFIAANC GSRRANNVTKSGKEQVDQIKDIREFEKNRVDRVVLTANTERSNVITGLNDTME NLFSLEKNESEISPSLTVAICVLENIFPINGSPOMTFVGILIDLAIRNCILIGDDI FKSQSDTLKSVLDPVLGAGIKPTSIENSYNHGNNGNLAPOTFSKETSKSNVVD DMVSNNALIVEPGHPHVIVIKVPPVYGDKRAMDEMSRIIPMGKNTIYLANTCEE SLAAPILLDIYTLAEIRSTQLCADGEKSHSPAVTII SYLTAKLVLPGRPVVAN ALSQRAMLENIILRACIGLSPEENNMLEYK"		
ORIGIN			
Query Match	2.6%; Score 45; DB 8; Length 1391;		
Best Local Similarity	100.0%; Pred. No. 3.7e-13;		
Matches	45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	953 TTTGATTGGTGAGATGACTTCAAGAGTGTGCAGACCATAATGAA	997	
Db	462 TTTCATTGGTGAAGATGACTTCAAGAGTGTGCAGACCATAATGAA	506	
RESULT 8			
AP006419/c			
LOCUS			
DEFINITION	Lotus corniculatus var. japonicus genomic DNA, chromosome 4, clone:LfJ43018, TM0307, complete sequence.		
ACCESSION	AP006419		
VERSION	AP006419.1 GI:31581050		
KEYWORDS	HTG.		
SOURCE	Lotus corniculatus var. japonicus (Lotus japonicus)		
ORGANISM	Lotus corniculatus var. japonicus Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoidae; Loteae; Lotus.		
REFERENCE			
AUTHORS	Asamizu,E., Kato,T., Sato,S., Nakamura,Y., Kaneko,T. and Tabata,S.		
TITLE	Structural Analysis of a Lotus japonicus Genome. IV. Sequence Features and Mapping of seventy-three TAC clones which cover the 7.5 Mb Regions of the Genome		

JOURNAL DNA Res. (2003) In press
REFERENCE 2 (bases 1 to 119858)
AUTHORS Sato, S.
TITLE Direct Submission
JOURNAL Submitted (07-MAY-2003) Shusui Sato, Kazusa DNA Research Institute, Department of Plant Gene Research; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail:satok@kazusa.or.jp, URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337), Fax:81-438-52-3934)

FEATURES
source location/Qualifiers
1..119858
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="4"
/clone="LjT43018"
/clone_lib="LjT library"
/note="TAC clone:TM0307-synonym: Lotus japonicus"

ORIGIN
Query Match 2.3%; Score 40; DB 8; Length 119858;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 1175 TGAGCTGTGTGACATCCAGACCATGTTGTTATTAAAG 1214
|||||
Db 46349 TGAGCTGTGTGACATCCAGACCATGTTGTTATTAAAG 46310
|||||

RESULT 9
CPINOIG 1978 bp DNA linear Pln 26-JUN-1995
LOCUS C.paradis1 (Macf) INO1 gene.
DEFINITION Z32632
ACCESSION Z32632.1 GI:602564
VERSION INO1 gene.
KEYWORDS Citrus x paradis1
ORGANISM Citrus x paradis1
SOURCE Citrus x paradis1
Bakayota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Sapindales; Rutaceae; Citrus.
1 (bases 1 to 1978)
Holland, D.
REFERENCE Direct Submission
AUTHORS Submitted (08-APR-1994) Doron Holland, Fruit Tree Breeding and
TITLE Genetics, Volcani Research, Organisation Center, Bet Dagan, 50250,
JOURNAL Israel

FEATURES
source location/Qualifiers
1..1978
/organism="Citrus x paradis1"
/mol_type="genomic DNA"
/db_xref="taxon:37656"
/clone="C38"
/rissue_type="fully expanded leaves"
/dev_stage="juvenile plants"
188..1711
/gene="INO1"
188..1711
/gene="INO1"
/codon_start=1
/protein_id="CAA83565.1"
/db_xref="GI:602565"
/db_xref="GDA:P42802"
/db_xref="UniProt/Swiss-Prot:P42802"
/translation="MFIENFKVESPVKTYDHIHSYYDYETTELVHENNGTYOMIV
KSTLVKBEFTDVAHVRLGVMALVWGNGNSSTLTGVIARBEICVATKTDVQDANF
GSLTQSAIRVDSVNGEIRIYAPKSKILPMWPPDIYFGWDIDDMNLAIAARVDF
IDLOKRLPYMESMVP.LPGIYDPDFLAAANGSANNVIKGTKEOMQI.IKDIFPKG
KNVDLVVYLVNTANTERSNVITYGLNDVYSLASLDKNAEATSPSLVIAIACLENI
PFINSGPONTFVGLDLARRNCILGDDPKSGQTKMSVLYVDFVGAIGKIPSTIYS
YNLQNDGNNLSAPQTFPSKEISKSNVVDVWSSNVFNGLVNTRPRWKVYPYVAI
ERMADRYTSLFMWGSKSTIVLHNTCEBSLLAAVLIIDLVLAELSTRIQKABEGSGF

ORIGIN	HSHFHPVATILSYLTKAPLVPGEVNVNALSQKRAMLENILIRACVGLAPENNMLLEYK"									
Query Match	2.0%	Score 35;	DB 8;	Length 1978;						
Best Local Similarity	100.0%;	Pred. No. 1.8e-07;								
Matches 35;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;						
Qy	246	ATGCTTGCGGTTGGGCGAACAACAGCGCTCAAC	280							
Db	380	ATGCTTGCGGTTGGGCGAACAACAGCGCTCAAC	414							
RESULT 10										
LOCUS	MCU32511	2053 bp		mRNA	linear	PLN 11-JUL-1996				
DEFINITION	Mesembryanthemum crystallinum myo-inositol-1-phosphate synthase									
ACCESSION	U32511									
VERSION	U32511									
KEYWORDS	U32511.1 GI:975887									
SOURCE	Mesembryanthemum crystallinum (common iceplant)									
ORGANISM	Mesembryanthemum crystallinum									
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Aizoaceae; Mesembryanthemum.									
AUTHORS	1 (bases 1 to 2053) Ishiteni,M., Majumder,A.L., Bornhouser,A., Michalowski,C.B., Jensen,R.G. and Bohnerl,H.J.									
TITLE	Coordinate transcriptional induction of myo-inositol metabolism during environmental stress									
JOURNAL	Plant J. 9 (4), 537-548 (1996)									
MEDLINE	96208959									
PUBMED	8624516									
REFERENCE	2 (bases 1 to 2053) Ishiteni,M., Majumder,A.L., Bornhouser,A., Michalowski,C.B., Jensen,R.G. and Bohnerl,H.J.									
AUTHORS	Direct Submission									
TITLE	Submitted (27-JUL-1995) Hans J. Bohnerl, Biochemistry, University of Arizona, BioSciences West 516, Tucson, AZ 85721, USA									
JOURNAL	Location/Qualifiers									
FEATURES	1..2053									
source	/organism="Mesembryanthemum crystallinum"									
	/mol_type="mRNA"									
	/db_xref="taxon:3544"									
	182..1720									
CDS	/codon_start=1									
	/product="myo-inositol-1-phosphate synthase"									
	/protein_id="AAB03687.1"									
	/db_xref="GI:975888"									
	translation="MFPLSEFVSEPNRYKTENEISVYNYDTTFLVHNRKDAQGYQMIVKRTVQVYHKTDTRVRLKGLVMVGWGNNGSTLTGCVIANRGISWATDKIQANAYFGSLTQASLRVGSFNGBEITYAFPKSLPWNPDVDFGMDISDMADAMTRAVFEDILOLQRPYEMHMYPLPIYDPDLAQAQSRANVIGTKKEQVERVTKDIREFEKKRRKVDVVVILMTGNTERYSNVVAVGDDMLDNLIALEKKESEISPSYALACIEEINIPINGSRQNTFVPLGLDILAKNSLIGDQDFRSGCQKMSYVDPFLVAGICPSTISVYHLGNNDGNNLSAPOTFRSKISXSNVVDVVASGIIYEGEHPDHVYVTKTYVYVGSKRAMDEYTSIEIFMGGNTITVMHNTCDSSLAAPIIIDLVLAELSTRIDLKAEEDDFHSHFHPVATILSYLTKAPLVPGEVNVNALSQKRAMDENILIRACVGLAPENNMILEYK"									
ORIGIN										
Query Match	1.8%	Score 32;	DB 8;	Length 2053;						
Best Local Similarity	100.0%;	Pred. No. 9.2e-06;								
Matches 32;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;						
Qy	1233	GACAGCAAGAGCGATGATGACGACACTTC	1264							
Db	1367	GACAGCAAGAGCGATGATGACGACACTTC	1398							
RESULT 11										
LOCUS	AC125389/c	143012 bp		DNA	linear	PLN 15-JUN-2004				
	AC125389									

DEFINITION	Medicago truncatula clone mth2-12a18, complete sequence.			
ACCESSION	AC125389			
VERSION	AC125389.32 GI:48717535			
KEYWORDS	HTG.			
SOURCE	Medicago truncatula (barrel medic)			
ORGANISM	Medicago truncatula			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eutroside I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.			
AUTHORS	1 (bases 1 to 143012)			
TITLE	Shauli,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.			
REFERENCE	Medicago truncatula BAC Clone mth2-12a18			
AUTHORS	2 (bases 1 to 143012)			
TITLE	Shauli,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.			
JOURNAL	Direct Submission			
REFERENCE	Submitted (25-JUN-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
AUTHORS	3 (bases 1 to 143012)			
TITLE	Shauli,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.			
JOURNAL	Direct Submission			
REFERENCE	Submitted (20-MAY-2004) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
AUTHORS	4 (bases 1 to 143012)			
TITLE	Shauli,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.			
JOURNAL	Direct Submission			
REFERENCE	Submitted (15-JUN-2004) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
COMMENT	On Jun 15, 2004 this sequence version replaced gi:47524030.			
FEATURES	-----			
SOURCE	Location/Qualifiers			
	1..143012			
	/organism="Medicago truncatula"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:3880"			
	/clone="mth2-12a18"			
ORIGIN	/clone_1fb="Medicago truncatula BAC library H2"			

Query Match	1.8%	Score 32;	DB 8;	Length 143012;
Best Local Similarity	100.0%;	Pred. No. 7.2e-06;		
Matches 32;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1479	ACACCACTGGTGATGCATTGTCAACACAGG	1510	
Db	13378	ACACCACTGGTGATGCATTGTCAACACAGG	13347	
RESULT 12				
LOCUS	AV323824	1533 bp	mRNA	linear
DEFINITION	Xerophyta viscosa myo-inositol-1-phosphate synthase INOI mRNA, complete cds.			
ACCESSION	AV323824			
VERSION	AV323824.1	GI:32492881		
KEYWORDS	Xerophyta viscosa			
SOURCE	Xerophyta viscosa			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Velloziaceae; Xerophyta.			
REFERENCE	1 (bases 1 to 1533)			

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:47:21 ; Search time 966.851 Seconds
(without alignments)
10775.958 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctcttcttcctcttgcgt.....aatgtatcataatttgcgt 1760

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1760	100.0	1760	12	ADQ14490	Adq14490 Wild type
2	1760	100.0	1760	2	AAV62440	AAV62440 Soybean w
3	1533	87.1	1533	12	ADQ14498	Adq14498 Mutant so
4	1533	87.1	1533	13	AD881999	Ad881999 Soybean m
5	1533	87.1	1533	2	AAV62443	AAV62443 Soybean m
6	1482	84.2	1533	12	ADQ14494	Adq14494 Mutant so
7	1482	84.2	1533	13	AD881997	Ad881997 Soybean m
8	1482	84.2	1533	12	ADQ14504	Adq14504 Wild type
9	146	8.3	1533	12	ADQ14500	Adq14500 Mutant so
10	146	8.3	1533	12	ADQ14502	Adq14502 Mutant so
11	146	8.3	1533	13	AD882005	Ad882005 Soybean m
12	146	8.3	1533	13	AD882003	Ad882003 Soybean m
13	146	8.3	1533	13	AD882001	Ad882001 Soybean m
14	146	8.3	1533	13	AD882001	Ad882001 Soybean m
15	29	1.6	377	12	ADP93370	Adp93370 Cotton ex
16	29	1.6	566	13	ACN57874	Acn57874 Cotton gy
17	29	1.6	618	13	ACN62990	Acn62990 Cotton ca
18	26	1.5	1781	4	AAV62441	AAV62441 Soybean m
19	24	1.4	35	2	AAV62441	AAV62441 Soybean m
20	24	1.4	35	12	ADQ14492	Adq14492 Soybean m

21	23	1.3	345	12	ADP94497	Adp94497 Cotton ex
22	23	1.3	457	12	ADJ10863	Adj10863 Recombina
23	23	1.3	557	13	ACN58059	Acn58059 Cotton gy
24	23	1.3	601	13	ACN60643	Acn60643 Cotton gy
25	23	1.3	605	13	ACN60421	Acn60421 Cotton gy
26	23	1.3	1536	12	ADP43918	Adp43918 P. coarct
27	23	1.3	1578	2	AAV91656	AAV91656 Inositol-
28	23	1.3	1578	2	AAV65400	AAV65400 Inositol-
29	23	1.3	1665	2	AAV90006	AAV90006 Inducible
30	23	1.3	1931	2	AAV24407	AAV24407 Maize myo
31	23	1.3	1950	2	AAV90402	AAV90402 Nicotiana
32	23	1.3	3546	2	AAV24410	AAV24410 Maize myo
33	23	1.3	3546	2	AAV24411	AAV24411 Maize myo
34	21	1.2	35	13	AD881995	Ad881995 PCR prime
35	21	1.2	39	12	AAV62442	AAV62442 Soybean m
36	21	1.2	39	12	ADQ14493	Adq14493 Soybean m
37	20	1.1	257	12	ADP95252	Adp95252 Cotton ex
38	20	1.1	291	12	ADP93783	Adp93783 Cotton ex
39	20	1.1	388	13	ACN52468	Acn52468 Cotton an
40	20	1.1	389	13	ACN51365	Acn51365 Cotton an
41	20	1.1	449	12	ADP91072	Adp91072 Cotton ex
42	20	1.1	506	13	ACN57963	Acn57963 Cotton gy
43	20	1.1	512	10	AD881407	Ad881407 Arabidops
44	20	1.1	517	13	ACN63079	Acn63079 Cotton ca
45	20	1.1	525	13	ACN58143	Acn58143 Cotton gy

ALIGNMENTS

RESULT 1	ADQ14490	standard; cDNA, 1760 BP.
ID	ADQ14490	
XX	ADQ14490;	
AC		
XX		
DT	23-SEP-2004	(first entry)
XX		
DE	Wild type soybean myo-inositol 1-phosphate synthase cDNA #1.	
XX		
XX	Soybean; myo-inositol 1-phosphate synthase; gene; ss;	
KW	myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;	
KM	inorganic phosphate.	
XX		
OS	Glycine max.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	54..1586
FT		/tag= a
FT		/product= "Wild type soybean myo-inositol 1-phosphate synthase #1"
XX		
PN	US2004128713-A1.	
XX		
PD	01-JUL-2004.	
XX		
PF	21-NOV-2003; 2003US-00718952.	
XX		
PR	08-APR-1997; 97US-00835751.	
PR	07-APR-1998; 98WO-US006822.	
PR	26-APR-1999; 99US-00299315.	
PR	11-MAR-2002; 2002US-00025003.	
XX		
PA	(HITZ/) HITZ W D.	
PA	(SEBA/) SEBASTIAN S A.	
PA	(GRAC/) GRACE D J.	
PA	(STRE/) STREIT L G.	
XX		
PI	Hitz WD, Sebastian SA, Grace DJ, Streit LG;	
XX		
DR	WPI; 2004-533135/51.	
DR	P-PsDB; ADQ14491.	
XX		

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

XX Claim 4; SEQ ID NO 1; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of lipo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1760 BP; 494 A; 371 C; 411 G; 484 T; 0 U; 0 Other;

Query Match 100.0%; Score 1760; DB 12; Length 1760;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCTTTTATTCCTTTGTAATTCATTCATTCTTATCTTTGTAATAATATGTGA 60
DB 1 CTTCTTTTATTCCTTTGTAATTCATTCATTCTTATCTTTGTAATAATATGTGA 60
QY 61 TCGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCCGAGCTGATTCAGTCCG 120
DB 61 TCGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCCGAGCTGATTCAGTCCG 120
QY 121 TGTACAACCTACGAAACCAACCTTGTTCACGAGAACGAAATGGCACTTACGTGA 180
DB 121 TGTACAACCTACGAAACCAACCTTGTTCACGAGAACGAAATGGCACTTACGTGA 180
QY 181 TTGTCAACCCCAATGTGTCAAAATTAAGAAATTTAAACCAACATCCATGTTCCAAATTG 240
DB 181 TTGTCAACCCCAATGTGTCAAAATTAAGAAATTTAAACCAACATCCATGTTCCAAATTG 240
QY 241 GGGTAAATGCTTGGGTGGGTGGAAACAACGGCTCAACCCCTCACGGGTGTTATTG 300
DB 241 GGGTAAATGCTTGGGTGGGTGGAAACAACGGCTCAACCCCTCACGGGTGTTATTG 300
QY 301 CTAAACGAGGGGCAATTTCAATGGCTTACAAAGACAAATTCATTTG 360
DB 301 CTAAACGAGGGGCAATTTCAATGGCTTACAAAGACAAATTCATTTG 360
QY 361 GGTCCCTCAACCAAGGCTCAGCTATCCGAGTTGGGTCTTCCAGGGAGAGAAATCTATG 420
DB 361 GGTCCCTCAACCAAGGCTCAGCTATCCGAGTTGGGTCTTCCAGGGAGAGAAATCTATG 420
QY 421 CCCCATTTCAAGAGCTGTTCCTCAATGTTAAACCTGACGACATTTGTTGGGGATGGG 480
DB 421 CCCCATTTCAAGAGCTGTTCCTCAATGTTAAACCTGACGACATTTGTTGGGGATGGG 480
QY 481 ATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTGTTGACATGATT 540
DB 481 ATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTGTTGACATGATT 540
QY 541 TGCAGAAAGAGTTGAGGCTTACATGGAATCCATGCTTCCATCCCGGAAATTAAGACC 600
DB 541 TGCAGAAAGAGTTGAGGCTTACATGGAATCCATGCTTCCATCCCGGAAATTAAGACC 600
QY 601 CGGATTTTCATCTGTCCAAACCAAGAGAGCGTGCACAAACGTTCATCAAGGGCACAAGC 660

DB 601 CGGATTTTCATCTGTCCAAACCAAGAGAGCGTGCACAAACGTTCATCAAGGGCACAAGC 660
QY 661 AAGACCAAGTTCAAAATTCATCAAGACATCAAGGCGTTTAAAGAGCCACAAGTGG 720
DB 661 AAGACCAAGTTCAAAATTCATCAAGACATCAAGGCGTTTAAAGAGCCACAAGTGG 720
QY 721 ACAAGTGGTGTACTGTGACTGCACTGCCAACAGAGAGGTACATTAATTTGTTGGGCC 780
DB 721 ACAAGTGGTGTACTGTGACTGCACTGCCAACAGAGAGGTACATTAATTTGTTGGGCC 780
QY 781 TTAATGACACCATGAGAAATCTTTGGCTGCTGTGACAGAAATAGAGCTGAGATTTCTC 840
DB 781 TTAATGACACCATGAGAAATCTTTGGCTGCTGTGACAGAAATAGAGCTGAGATTTCTC 840
QY 841 CTTCAACCTGTATACCCATGCTGTGTATGGAATAATGTTCTTCAATTAAGAGCC 900
DB 841 CTTCAACCTGTGTATACCCATGCTGTGTATGGAATAATGTTCTTCAATTAAGAGCC 900
QY 901 CTCAAGACATTTTGTACAGGGCTGATTTGATCTTGCATCGCAGAGAACACTTTGATTG 960
DB 901 CTCAAGACATTTTGTACAGGGCTGATTTGATCTTGCATCGCAGAGAACACTTTGATTG 960
QY 961 GTGAGATGACTTCAAGAGTGTGACACCAGAAATGAATCTGTGTTGGTTCTTTC 1020
DB 961 GTGAGATGACTTCAAGAGTGTGACACCAGAAATGAATCTGTGTTGGTTCTTTC 1020
QY 1021 TGGGGGCTGTATCAAGCCAACTATATGATGATTAACCATCTGGGAAACAATGATG 1080
DB 1021 TGGGGGCTGTATCAAGCCAACTATATGATGATTAACCATCTGGGAAACAATGATG 1080
QY 1081 GTATGAATCTTTGGGCTCCACAACTTTCGTTCCAGAAATCTCCAGAGCAACGTTG 1140
DB 1081 GTATGAATCTTTGGGCTCCACAACTTTCGTTCCAGAAATCTCCAGAGCAACGTTG 1140
QY 1141 TTGATGATATGTTCAACAGCAATGCCATCTCTATGAGCTGTGTAAATCCAGCATG 1200
DB 1141 TTGATGATATGTTCAACAGCAATGCCATCTCTATGAGCTGTGTAAATCCAGCATG 1200
QY 1201 TTGTTGTTAATTAAGTATGTCCTTAAGTGGGACAGCAAGAGCACTGATGATGTA 1260
DB 1201 TTGTTGTTAATTAAGTATGTCCTTAAGTGGGACAGCAAGAGCACTGATGATGTA 1260
QY 1261 CTTCAAGATATTCATGGGTGGAAAGACCACTTTTGGACAACAATCCAGAGATT 1320
DB 1261 CTTCAAGATATTCATGGGTGGAAAGACCACTTTTGGACAACAATCCAGAGATT 1320
QY 1321 CCTCTTAAGCTGCTCTATTAATCTTGAATGTGCTCTTCTGTAAGCTCAGACTAGAA 1380
DB 1321 CCTCTTAAGCTGCTCTATTAATCTTGAATGTGCTCTTCTGTAAGCTCAGACTAGAA 1380
QY 1381 TCGAGTTTAAACCTGAAAATGAGGAAAATTCACATCTCCACCAAGTGTATCAATCC 1440
DB 1381 TCGAGTTTAAACCTGAAAATGAGGAAAATTCACATCTCCACCAAGTGTATCAATCC 1440
QY 1441 TCAGTACTCTCAACAGGCTCCTGTGTTCAACCGGGTACACAGTGTGATGATTGT 1500
DB 1441 TCAGTACTCTCAACAGGCTCCTGTGTTCAACCGGGTACACAGTGTGATGATTGT 1500
QY 1501 CAAAGCAGCTGCAATGCTGAAAAATATAGAGGCTTTGTGTGATTTGGCCCAAGAGA 1560
DB 1501 CAAAGCAGCTGCAATGCTGAAAAATATAGAGGCTTTGTGTGATTTGGCCCAAGAGA 1560
QY 1561 ATTAACATGATTCGAGATCAAAAGTAAAGATGGGACCGAAGAAATTAATGTTGGGGT 1620
DB 1561 ATTAACATGATTCGAGATCAAAAGTAAAGATGGGACCGAAGAAATTAATGTTGGGGT 1620
QY 1621 CCTAGCTGATTTTATGTTAATATATGTTGCTTAATTAATTTGCAAGTGAATTGAA 1680
DB 1621 CCTAGCTGATTTTATGTTAATATATGTTGCTTAATTAATTTGCAAGTGAATTGAA 1680
QY 1681 TGCATCAGCTTATTAATGCTTTAAGCGGGCATATTCGTTTACATGAGAACTGATG 1740


```

QY 1141 TTGATGATATGTCACAGCAATGCAATGCTCTTATGAGCCTGCTGAACATCCAGACCATG 1200
DB 1141 TTGATGATATGTCACAGCAATGCAATGCTCTTATGAGCCTGCTGAACATCCAGACCATG 1200
QY 1201 TTGTTGTTATTAAGTATGTCCTTACGTAAGGAGGACAGCAAGAGGACCATGATGATTA 1260
DB 1201 TTGTTGTTATTAAGTATGTCCTTACGTAAGGAGGACAGCAAGAGGACCATGATGATTA 1260
QY 1261 CTTGAGAGATATTCATGAGGTGAGAAAGAGACCATTTGTTTGACAAACATCCAGAGATT 1320
DB 1261 CTTGAGAGATATTCATGAGGTGAGAAAGAGACCATTTGTTTGACAAACATCCAGAGATT 1320
QY 1321 CCTCTTGAGCTGCTCTTATTAATCTTGAGACTTGTCTTCTTGCTGAGCTCAGCACTAAGA 1380
DB 1321 CCTCTTGAGCTGCTCTTATTAATCTTGAGACTTGTCTTCTTGCTGAGCTCAGCACTAAGA 1380
QY 1381 TCGAGTTTAAAGCTGAGAAATGAGGAGAAATTCATCTCCAGCCAGTTGCTACCATCC 1440
DB 1381 TCGAGTTTAAAGCTGAGAAATGAGGAGAAATTCATCTCCAGCCAGTTGCTACCATCC 1440
QY 1441 TCGAGTTTAAAGCTGAGAAATGAGGAGAAATTCATCTCCAGCCAGTTGCTACCATCC 1500
DB 1441 TCGAGTTTAAAGCTGAGAAATGAGGAGAAATTCATCTCCAGCCAGTTGCTACCATCC 1500
QY 1501 CAAAGCAGCGCTGCAATGCTGAGAAACATTAATGAGGAGCTTGTGTTGATTTGCCCCAGAGA 1560
DB 1501 CAAAGCAGCGCTGCAATGCTGAGAAACATTAATGAGGAGCTTGTGTTGATTTGCCCCAGAGA 1560
QY 1561 ATAAATGATTTCTGAGTACAAAGTGAAGCATGAGGACGGAACAAATTAATTAATTTGGGTAG 1620
DB 1561 ATAAATGATTTCTGAGTACAAAGTGAAGCATGAGGACGGAACAAATTAATTAATTTGGGTAG 1620
QY 1621 CTAAGCTGAATGTTTATGTTAATTAATTAATGTTTGAATTTGCAAGTGAATTTGAA 1680
DB 1621 CTAAGCTGAATGTTTATGTTAATTAATTAATGTTTGAATTTGCAAGTGAATTTGAA 1680
QY 1681 TGATGAGCTTCATTAATGCTTTAGACGGGGGCAATTTGCTTTAGTAGGAACATGATG 1740
DB 1681 TGATGAGCTTCATTAATGCTTTAGACGGGGGCAATTTGCTTTAGTAGGAACATGATG 1740
QY 1741 AATGATGATTAATTTGTTGT 1760
DB 1741 AATGATGATTAATTTGTTGT 1760

```

```

RESULT 3
ADQ14498
ID ADQ14498 standard; cDNA; 1533 BP.
XX
AC ADQ14498;

```

```

23-SEP-2004 (first entry)

```

```

Mutant soybean myo-inositol 1-phosphate synthase cDNA #2.

```

```

DE Soybean, myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate; mutant.
OS Glycine max.
OS Synthetic.
XX

```

```

FH Key Location/Qualifiers
FT 1..1533
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT /tag= a
FT /synthase #2"

```

```

US2004128713-A1.

```

```

01-JUL-2004.

```

```

PF 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98MO-US006822.
PR 26-APR-1999; 99US-0029315.
PR 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ// HITZ W D.
PA (SEBA// SEBASTIAN S A.
PA (GRAC// GRACE D J.
PA (STRE// STREIT L G.
PI
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
DR WPI; 2004-533135/51.
DR P-PSDB; ADQ14499.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX Example 8; SEQ ID NO 9; 48pp; English.
PS
PS The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
XX Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;
SQ

```

```

Query Match 87.1%; Score 1533; DB 12; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 54 ATGTTCAATCGAGATTTTAAAGTTGATGTCCTTAATGTAAGTACACCGAGCTGAGATT 113
DB 1 ATGTTCAATCGAGATTTTAAAGTTGATGTCCTTAATGTAAGTACACCGAGCTGAGATT 60
QY 114 CAGTCCGCTGATCAATTAAGAAACCAACGAACTTTGTCACGAGACAGAAATGCACTAT 173
DB 61 CAGTCCGCTGATCAATTAAGAAACCAACGAACTTTGTCACGAGACAGAAATGCACTAT 120
QY 174 CAGTGGATTGCAAAACCAAAATGCTCAAAATTAAGAAATTAACCAACATCATGTTCTT 233
DB 121 CAGTGGATTGCAAAACCAAAATGCTCAAAATTAAGAAATTAACCAACATCATGTTCTT 180
QY 234 AAATTAAGGGTAAATGCTTGTGGGTGGAACCAACGAGCTCAACCTCAACGGGTGT 293
DB 181 AAATTAAGGGTAAATGCTTGTGGGTGGAACCAACGAGCTCAACCTCAACGGGTGT 240
QY 294 GTTATGTTAAACGAGAGGGGATTTCAATGGGCTCAAAAGACCAAGATTCAACAGCCAT 353
DB 241 GTTATGTTAAACGAGAGGGGATTTCAATGGGCTCAAAAGACCAAGATTCAACAGCCAT 300
QY 354 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGGAA 413
DB 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGGAA 360

```

QY 414 ATCTATGCCCCCATCTCAAGAGCGCTGCTTCAATGGTTAAACCTGACGACATTTGTTGGG 473
 DB 361 ATCTATGCCCCCATCTCAAGAGCGCTGCTTCAATGGTTAAACCTGACGACATTTGTTGGG 420
 QY 474 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTTGTTGAC 533
 DB 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTTGTTGAC 480
 QY 534 ATGATATTTGCAGAGCAAGCTTGAAGGCTTACATGAAATCCATGCTTCCATCTCCCGGAAATC 593
 DB 481 ATGATATTTGCAGAGCAAGCTTGAAGGCTTACATGAAATCCATGCTTCCATCTCCCGGAAATC 540
 QY 594 TATGACCCGGATTTCAATGCTGCTCAACAGAGAGAGGCTGCCAACAACGTCATCAAGGGG 653
 DB 541 TATGACCCGGATTTCAATGCTGCTCAACAGAGAGAGGCTGCCAACAACGTCATCAAGGGG 600
 QY 654 ACAAGCAAGAGCAAGCTTCAACAAATCATCAAGAGCATCAAGGCGTTTAAAGAGCCACC 713
 DB 601 ACAAGCAAGAGCAAGCTTCAACAAATCATCAAGAGCATCAAGGCGTTTAAAGAGCCACC 660
 QY 714 AAAGTGCAAGAGGTTGTTGATCTGTGACCTGCCAACAAGAGGTTACGTAATTTGGTT 773
 DB 661 AAAGTGCAAGAGGTTGTTGATCTGTGACCTGCCAACAAGAGGTTACGTAATTTGGTT 720
 QY 774 GTGGGCGCTTAATGACCAATGAGAAATCTTGGCTGTGGTGGACAGAAATGAGGCTGAG 833
 DB 721 GTGGGCGCTTAATGACCAATGAGAAATCTTGGCTGTGGTGGACAGAAATGAGGCTGAG 780
 QY 834 ATTTCTCTTCCACCTTGTATGTCATTTGTTGTTATGGAATAATGTTCTTTTCAATTAAT 893
 DB 781 ATTTCTCTTCCACCTTGTATGTCATTTGTTGTTATGGAATAATGTTCTTTTCAATTAAT 840
 QY 894 GGAAGCCCTCAGAAACCTTTTGTACCAAGGCTGATTTGATTTTCCATCCGAGAGAACT 953
 DB 841 GGAAGCCCTCAGAAACCTTTTGTACCAAGGCTGATTTGATTTTCCATCCGAGAGAACT 900
 QY 954 TTGATTTGGTGAAGATGATCAAGAGTGGTCAAGCAAAATGAAATCTGTTGTTGAT 1013
 DB 901 TTGATTTGGTGAAGATGATCAAGAGTGGTCAAGCAAAATGAAATCTGTTGTTGAT 960
 QY 1014 TTCTTTTGGGGGCTGTATCAAGCCAACTTATAGTCAGTTTCAACCACTCTGGGAAAC 1073
 DB 961 TTCTTTTGGGGGCTGTATCAAGCCAACTTATAGTCAGTTTCAACCACTCTGGGAAAC 1020
 QY 1074 AATGATGATGATATTTTGGCTCCACAACTTTCCGTTCCAGAGAAATCTCCAAAGAC 1133
 DB 1021 AATGATGATGATATTTTGGCTCCACAACTTTCCGTTCCAGAGAAATCTCCAAAGAC 1080
 QY 1134 AACGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
 DB 1081 AACGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 QY 1194 GACCATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
 DB 1141 GACCATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 QY 1254 GAGTACACTTCAAGATATTCATGAGTGGTGAAGAGACCAATGTTTTCACCAACATGCT 1313
 DB 1201 GAGTACACTTCAAGATATTCATGAGTGGTGAAGAGACCAATGTTTTCACCAACATGCT 1260
 QY 1314 GAGGATTCCTCTTACTGCTCTATATTCCTTGGATGATGATGATGATGATGATGATGATGAT 1373
 DB 1261 GAGGATTCCTCTTACTGCTCTATATTCCTTGGATGATGATGATGATGATGATGATGATGAT 1320
 QY 1374 ACTAGATGAGATTTAAAGCTGAATAATGAGGAAATTCACATCATTTCCACCAAGTTGCT 1433
 DB 1321 ACTAGATGAGATTTAAAGCTGAATAATGAGGAAATTCACATCATTTCCACCAAGTTGCT 1380
 QY 1434 ACCATCTTCAAGCTTCAACAAAGGCTCTCTGTTTCCACCGGGTACCAACAGTGTGAAT 1493
 DB 1381 ACCATCTTCAAGCTTCAACAAAGGCTCTCTGTTTCCACCGGGTACCAACAGTGTGAAT 1440
 QY 1494 GCATTGTCAAGAGCGGTGCAATGCTGAAATAATAGAGGCTTGTGTGATTTGGGC 1553

DB 1441 GCATTGTCAAGAGCGGTGCAATGCTGAAATAATAGAGGCTTGTGTGATTTGGCC 1500
 QY 1554 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1586
 DB 1501 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1533
 RESULT 4
 ADS81999
 ID ADS81999 standard; cDNA; 1533 BP.
 XX
 AC ADS81999;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 3.
 XX
 KM Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
 XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
 XX
 OS Glycine max; line 29004JP01.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1533
 FT /tag= a
 FT /product= "myo-inositol 1-phosphate synthase"
 XX
 PN US2003074685-A1.
 PD 17-APR-2003.
 XX
 PF 11-MAR-2002; 2002US-00025003.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 XX
 PI Hitz WD, Sebastian SA;
 XX
 DR WPI: 2004-639957/62.
 DR P-PSDB; ADS82000.
 XX
 PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
 PT
 XX
 PS Example 8; SEQ ID NO 9; 34bp; English.
 XX
 CC The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis of myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g), provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a

CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.

XX Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Query Match 87.1%; Score 1533; DB 13; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 54 ATGTCATGAGAAATTTAAAGTTGAGTCTCTAATGTAAGTACCGAGACTGAGAT 113
DB 1 ATGTCATGAGAAATTTAAAGTTGAGTCTCTAATGTAAGTACCGAGACTGAGAT 60
QY 114 CAGTCGGTGTACAACTACGAAACCAAGCACTTGTTCAGAGAACAGAAATGGCACTAT 173
DB 61 CAGTCGGTGTACAACTACGAAACCAAGCACTTGTTCAGAGAACAGAAATGGCACTAT 120
QY 174 CAGTCGGTGTACAACTACGAAACCAAGCACTTGTTCAGAGAACAGAAATGGCACTAT 233
DB 121 CAGTCGGTGTACAACTACGAAACCAAGCACTTGTTCAGAGAACAGAAATGGCACTAT 180
QY 234 AAATTAAGGGGTAAATGTTGGTGGGTGGGAAACAGGCTCAACCTCACCGGTGGT 293
DB 181 AAATTAAGGGGTAAATGTTGGTGGGTGGGAAACAGGCTCAACCTCACCGGTGGT 240
QY 294 GTTATTTGCTAACCGAAGGGCAATTCATGGGCTACAAAGCAAGATTCAGAAAGCAAT 353
DB 241 GTTATTTGCTAACCGAAGGGCAATTCATGGGCTACAAAGCAAGATTCAGAAAGCAAT 300
QY 354 TACTTTGGCTCCCTCAACCCAAAGCTCAGCTATCCAGTTGGGTCTTCCAGGAGAGAA 413
DB 301 TACTTTGGCTCCCTCAACCCAAAGCTCAGCTATCCAGTTGGGTCTTCCAGGAGAGAA 360
QY 414 ATCTAATGCCCAATTCAGAGAGCCGCTTCCAAATGGTTAAACCTGACGACATTTGGTTGG 473
DB 361 ATCTAATGCCCAATTCAGAGAGCCGCTTCCAAATGGTTAAACCTGACGACATTTGGTTGG 420
QY 474 GGATGGGATATCAGCAATGAACCTGGCTGATGCCATGGCCAGGCAAAAGGTGTTGAC 533
DB 421 GGATGGGATATCAGCAATGAACCTGGCTGATGCCATGGCCAGGCAAAAGGTGTTGAC 480
QY 534 ATCGATTTGCAAGAGAGTTGAGGCTTTACATGAATCCATGCTTCCACTCCCGGAATC 593
DB 481 ATCGATTTGCAAGAGAGTTGAGGCTTTACATGAATCCATGCTTCCACTCCCGGAATC 540
QY 594 TATGACCCCGGATTTCTTGTCTGCCAACAAAGAGACGGGCCAACACGTCATCAAGGGC 653
DB 541 TATGACCCCGGATTTCTTGTCTGCCAACAAAGAGACGGGCCAACACGTCATCAAGGGC 600
QY 654 ACAAGCAAGAGAGCAATTCACAAATCATCAAGATCAAGGCGTTTAAAGAGGCAAC 713
DB 601 ACAAGCAAGAGAGCAATTCACAAATCATCAAGATCAAGGCGTTTAAAGAGGCAAC 660
QY 714 AAAGTGACAGAGGTGTTGTTACTGTGACTGCAACACAGAGGTACAGTAATTTGGTT 773
DB 661 AAAGTGACAGAGGTGTTGTTACTGTGACTGCAACACAGAGGTACAGTAATTTGGTT 720
QY 774 GTGGGCTTAATACACCATGAGGAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 833
DB 721 GTGGGCTTAATACACCATGAGGAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
QY 834 ATTTCCTTCCACCTTGTATGCAATGCTTGTGTTATGGAATAATCTCTTCAATTAAT 893
DB 781 ATTTCCTTCCACCTTGTATGCAATGCTTGTGTTATGGAATAATCTCTTCAATTAAT 840
QY 894 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTCCATCGGAGGAACACT 953
DB 841 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTCCATCGGAGGAACACT 900
```

```
DB 841 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTCCATCGGAGGAACACT 900
QY 954 TTGATTTGGTGAAGATGACTTCAAGAGTGTGACGACCAAAATGAATCTGTGTGGTTAT 1013
DB 901 TTGATTTGGTGAAGATGACTTCAAGAGTGTGACGACCAAAATGAATCTGTGTGGTTAT 960
QY 1014 TTCTTTGGGGGGCTGGTATCAAGCCAACTATATGTCAGTTACACCAATCTGGGAAAC 1073
DB 961 TTCTTTGGGGGGCTGGTATCAAGCCAACTATATGTCAGTTACACCAATCTGGGAAAC 1020
QY 1074 AATGATGGTATGAATCTTGGCTCCCAAACTTCCGTTCCAGGAATCTCCAAAGAC 1133
DB 1021 AATGATGGTATGAATCTTGGCTCCCAAACTTCCGTTCCAGGAATCTCCAAAGAC 1080
QY 1134 AACGTTGTTGATATGATATGTCACAGCAATGCATCTCTATGAGCCTGTGTAACATCCA 1193
DB 1081 AACGTTGTTGATATGATATGTCACAGCAATGCATCTCTATGAGCCTGTGTAACATCCA 1140
QY 1194 GACCATGTTGTTGTTATTAATGATGTGCTTACGTAGGGGACAGCAAGAGCCATGGAT 1253
DB 1141 GACCATGTTGTTGTTATTAATGATGTGCTTACGTAGGGGACAGCAAGAGCCATGGAT 1200
QY 1254 GAGTACACTTCAGAGATATTTACATGGGTGAAAAGACACATTTGTTGACAAACATGAC 1313
DB 1201 GAGTACACTTCAGAGATATTTACATGGGTGAAAAGACACATTTGTTGACAAACATGAC 1260
QY 1314 GAGATTCCTCTTCTTACTGCTCTCTATTAATCTTGAATCTTGATGCTGAGCTCAGC 1373
DB 1261 GAGATTCCTCTTCTTACTGCTCTCTATTAATCTTGAATCTTGATGCTGAGCTCAGC 1320
QY 1374 ACTAGATTCGAGTTTAAAGCTGAAAATGAGGAAAATTCACCTCATTCACCAAGTTGCT 1433
DB 1321 ACTAGATTCGAGTTTAAAGCTGAAAATGAGGAAAATTCACCTCATTCACCAAGTTGCT 1380
QY 1434 ACCATCTCAGTACTCCTACCAAGGCTCCTGTTTCCACCGGGTACACAGTGTGAT 1493
DB 1381 ACCATCTCAGTACTCCTACCAAGGCTCCTGTTTCCACCGGGTACACAGTGTGAT 1440
QY 1494 GCATTGTCAAAGCAGCTGCAATGCTGGAATAACATATGAGGCTTGTGTGATTTGGCC 1553
DB 1441 GCATTGTCAAAGCAGCTGCAATGCTGGAATAACATATGAGGCTTGTGTGATTTGGCC 1500
QY 1554 CCAGAGATTAACATGATTTCTGAGTACAAATGA 1586
DB 1501 CCAGAGATTAACATGATTTCTGAGTACAAATGA 1533
RESULT 5
ADS81993
ID ADS81993 standard; cDNA; 1533 BP.
XX
AC ADS81993;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 1.
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX Glycine max; cultivar Wye.
XX
OS
XX
FH Key Location/Qualifiers
FT 1..1533
FT CDS /tag= a
FT product= "myo-inositol 1-phosphate synthase"
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
```

PR	08-APR-1997;	97US-00835751..
PR	07-APR-1998;	98MO-US006822..
XX		
PA	(HITZ/) HITZ W D.	
PA	(SEBA/) SEBASTIAN S A.	
XX		
PI	Hitz WD, Sebastian SA;	
XX		
DR	WPI: 2004-639957/62.	
DR	P-PSDB; A0581994.	
XX		
PT	Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-	
PT	phosphate synthase, useful for altering raffinose saccharide, sucrose,	
PT	phytic acid and inorganic phosphate content of soybean seeds.	
XX		
PS	Claim 2; SEQ ID NO 1; 34bp; English.	
XX		
CC	The invention relates to an isolated nucleic acid fragment encoding a	
CC	soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-	
CC	phosphate synthase having decreasing capacity for the synthesis for myo-	
CC	inositol-1-phosphate. Also included are a chimeric gene (comprising the	
CC	nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its	
CC	complement, subfragment or the complement of the subfragment, operably	
CC	linked to suitable regulatory sequences, where expression of the chimeric	
CC	gene results in a decrease in expression of an endogenous or native gene	
CC	encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant	
CC	comprising the chimeric gene (with a heritable phenotype of a seed phytic	
CC	acid content of less than 17 micromol/g, a seed content of raffinose plus	
CC	stachyose of less than 14.5 micromol/g, and a seed sucrose content of	
CC	greater than 200 micromol/g, provided that the plant is not LR33), seeds	
CC	from the plant, making a soybean plant with the heritable phenotype	
CC	(comprising crossing LR33 or the plant comprising the chimeric gene with	
CC	an elite soybean plant and selecting a progeny plant of the cross of	
CC	crossing step that has a heritable phenotype as mentioned above), seeds	
CC	of soybean plant made by the above method, a soy protein product derived	
CC	from seeds of a soybean plant (homozygous for one or more gene encoding a	
CC	mutant myo-inositol 1-phosphate synthase having decreased capacity for	
CC	the synthesis of myo-inositol 1-phosphate, where the gene confers a	
CC	heritable phenotype as mentioned above), and making or producing a	
CC	soybean protein product derived from seeds of a soybean plant with	
CC	heritable phenotype as mentioned above. The nucleic acid is useful for	
CC	altering raffinose saccharide, sucrose, phytic acid and inorganic	
CC	phosphate content of soybean seeds thus leading to valuable and useful	
CC	soybean products, since the presence of high concentration of raffinose	
CC	oligosaccharides in soy plants (and other legumes) can lead to flatulence	
CC	when consumed by humans. The present sequence encodes a wild-type myo-	
CC	inositol 1-phosphate synthase.	
XX		
SO	Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;	
QY	Query Match	87.1%; Score 1533; DB 13; Length 1533;
Db	Best Local Similarity 100.0%; Pred. No. 0;	
Db	Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	54 ATGTTTCATCGAAGATTTTAAGTTGAGTGTCTTAATGTGAATACCGAGACTGAGATT	113
Db	1 ATGTTTCATCGAAGATTTTAAGTTGAGTGTCTTAATGTGAATACCGAGACTGAGATT	60
QY	114 CAGTCCGGTGTACAAATACGAAACCCGAACTTTGTCACGAGAACGAGATGGCACTTAT	173
Db	61 CAGTCCGGTGTACAACTACGAAACCCGAACTTTGTCACGAGAACGAGATGGCACTTAT	120
QY	174 CAGTGGATGTTCACAAACCCAAATCTGTCAAAATTAACCAATCACTCATGTTCTCT	233
Db	121 CAGTGGATGTTCACAAACCCAAATCTGTCAAAATTAACCAATCACTCATGTTCTCT	180
QY	234 AAATTTAGGGGTAAATGCTTGTGGGTGTGGGTGGAACCAACGGCTCAACCTCAACGGTGT	293
Db	181 AAATTTAGGGGTAAATGCTTGTGGGTGTGGGTGGAACCAACGGCTCAACCTCAACGGTGT	240
QY	294 GTTATTTGCTAACCGAGAGGGCATTTTCATGTGGGCTCAAAAGACACAGATTTCACAGCCAAAT	353
Db	241 GTTATTTGCTAACCGAGAGGGCATTTTCATGTGGGCTCAAAAGACACAGATTTCACAGCCAAAT	300

QY	354	TACCTTGGCTCCCTACCCAAAGCCTCAAGCTATCCGAGTGGGGTCCCTTCACAGGAGAAGAA	413
Db	301	TACTTTGGCTCCCTACCCAAAGCCTCAAGCTATCCGAGTGGGGTCCCTTCACAGGAGAAGAA	360
QY	414	ATCTATGCGCCATTCAAGAGCCTCTTCCAAATGGTTAAACCTGTACGACATTTGTGTTGGG	473
Db	361	ATCTATGCGCCATTCAAGAGCCTCTTCCAAATGGTTAAACCTGTACGACATTTGTGTTGGG	420
QY	474	GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGCGCAAGAGTGTGTTGAC	533
Db	421	GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGCGCAAGAGTGTGTTGAC	480
QY	534	ATCGATTTGCAAGAAAGCACTTGAAGCCTTACATGGAATTCATGCTTTCCATCCCCGGGAATC	593
Db	481	ATCGATTTGCAAGAAAGCACTTGAAGCCTTACATGGAATTCATGCTTTCCATCCCCGGGAATC	540
QY	594	TATGACCCGGATTTCATTGCTGGCCAAACAAAGAGAGCGTGCCAAACACGTATCAAGAGGC	653
Db	541	TATGACCCGGATTTCATTGCTGGCCAAACAAAGAGAGCGTGCCAAACACGTATCAAGAGGC	600
QY	654	ACAAAGCAAGAGCAAGTTCAACAAATCATCAAGACAGCGGTTTAAAGAAAGCCACC	713
Db	601	ACAAAGCAAGAGCAAGTTCAACAAATCATCAAGACAGCGGTTTAAAGAAAGCCACC	660
QY	714	AAAGTGCACAAAGGTGGTTGTACTGTGGACTGCGCAACACAGAGAGTACGTAATTTGGTT	773
Db	661	AAAGTGCACAAAGGTGGTTGTACTGTGGACTGCGCAACACAGAGAGTACGTAATTTGGTT	720
QY	774	GTGGGCGCTTAATGACACCAATGGAATAATCTCTGGGCTGCTGTGGAACAGAAATGAGGCTGAG	833
Db	721	GTGGGCGCTTAATGACACCAATGGAATAATCTCTGGGCTGCTGTGGAACAGAAATGAGGCTGAG	780
QY	834	ATTTCTCCCTTCCACTTGTATGACATTCCTGTGTATATGAAATATGTTCTTTCAATTAT	893
Db	781	ATTTCTCCCTTCCACTTGTATGACATTCCTGTGTATATGAAATATGTTCTTTCAATTAT	840
QY	894	GGAAGCCCTCAGAACACTTTTGTACCAAGGGCTGATGTAATCTTGGCATTCGAGAGAAACCT	953
Db	841	GGAAGCCCTCAGAACACTTTTGTACCAAGGGCTGATGTAATCTTGGCATTCGAGAGAAACCT	900
QY	954	TTGATTTGTGAGATGATCTTCAAGAGTGCAGCCAAAATGAATCTGTGTGGTAT	1013
Db	901	TTGATTTGTGAGATGATCTTCAAGAGTGCAGCCAAAATGAATCTGTGTGGTAT	960
QY	1014	TTCTTTGCGGGGCTGGATTCAGAACCCAAACATCTATAGCTTTCAACCATCTGGGAAAC	1073
Db	961	TTCTTTGCGGGGCTGGATTCAGAACCCAAACATCTATAGCTTTCAACCATCTGGGAAAC	1020
QY	1074	AATGATGGTATGAAATCTTTCGGCTCCAACAACTTTCGGTTCCAGAGAAATCTCCAAAGAC	1133
Db	1021	AATGATGGTATGAAATCTTTCGGCTCCAACAACTTTCGGTTCCAGAGAAATCTCCAAAGAC	1080
QY	1134	AACGTTGTTGATGATATGATGTCACAGCAATGCCATCTATAGCCTGTGGAACATCCA	1193
Db	1081	AACGTTGTTGATGATATGATGTCACAGCAATGCCATCTATAGCCTGTGGAACATCCA	1140
QY	1194	GACCATGTTGTTGATTAAGTATGTCCTTAAGTAAAGGAGGACACAAAGAGGCAATGGAT	1253
Db	1141	GACCATGTTGTTGATTAAGTATGTCCTTAAGTAAAGGAGGACACAAAGAGGCAATGGAT	1200
QY	1254	GAGTACACTTCAGAGATTAATTCATGGGTGAAAGAGCAACATTTGTTTGCACAAACATATGC	1313
Db	1201	GAGTACACTTCAGAGATTAATTCATGGGTGAAAGAGCAACATTTGTTTGCACAAACATATGC	1266
QY	1314	GAGGATTCCTCTTAACTGCTCCTATATATCTTGAATTTGGACTTGGTCTTTCTTGGCTGAGCTACG	1373
Db	1261	GAGGATTCCTCTTAACTGCTCCTATATATCTTGAATTTGGACTTGGTCTTTCTTGGCTGAGCTACG	1320
QY	1374	ACTAGAAATGAGTTTAAAGCTGAAAGAAAGAGGAGAAATTCACATCATTCACCCAGTTGCT	1433
Db	1321	ACTAGAAATGAGTTTAAAGCTGAAAGAAAGAGGAGAAATTCACATCATTCACCCAGTTGCT	1380

QY 1434 ACCATCTCAGTACTACCTACCAAGGCTCCTCTGTTCCACCGGTAACACAGTGTGAT 1493
| | | | |
Db 1381 ACCATCTCAGTACTACCTACCAAGGCTCCTCTGTTCCACCGGTAACACAGTGTGAT 1440
| | | | |
QY 1494 GCATTGTCAAAGCAGGTGCAATGCTGAAAACATATATAGAGGCTTGTGTGATTGGCC 1553
| | | | |
Db 1441 GCATTGTCAAAGCAGGTGCAATGCTGAAAACATATATAGAGGCTTGTGTGATTGGCC 1500
| | | | |
QY 1554 CCAGAGATATACATGATTTCTCGAGTACAGTGA 1586
| | | | |
Db 1501 CCAGAGATATACATGATTTCTCGAGTACAGTGA 1533
| | | | |

RESULT 6
AAV62443
ID AAV62443 standard; cDNA; 1533 BP.

AAV62443;

17-OCT-2003 (revised)
02-FEB-1999 (first entry)

Soybean mutant myo-inositol 1-phosphate synthase cDNA.

Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
phytic acid; de.

Glycine max; line LR33.

W09845448-A1.

15-OCT-1998.

07-APR-1998: 98WO-US006822.

08-APR-1997: 97US-00835751.

(DUPO) DU PONT DE NEMOURS & CO E. I.

Hitz WD, Sebastian SA;

WPI; 1998-568353/48.

P-PSDB; AAW79741.

Soybean plants containing altered myo-inositol-1-phosphate gene - useful
for generating plants with altered levels of e.g. raffinose, stachyose,
phytic acid, etc.

Example 5; Page 48-49; 63p; English.

This is the nucleotide sequence of cDNA encoding a mutant soybean myo-
inositol 1-phosphate synthase (MI 1-PS) present in clone LR33-10. This
clone was isolated from a cDNA library of soybean line LR33 by PCR
amplification (see AAV62441-42). Line LR33 was obtained by chemical
mutagenesis of wild-type soybean genome and as a reduced raffinose
saccharide phenotype. MI 1-PS is involved in glucose metabolism to phytic
acid, raffinose and stachyose. Sequencing revealed a single base change
mutation (G to T at base 1241) in the LR33 sequence when compared to the
wild-type sequence (see AAV62440). The mutation results in a seed
phenotype of very low raffinose saccharide sugars, very high sucrose and
low phytic acid. The mutated MI 1-PS nucleic acid is used to alter the
raffinose saccharide, sucrose, phytic acid and inorganic phosphate
content of soybean seeds, leading to useful soybean products, e.g. a seed
phytic acid content of less than 17 ug/g, a seed content of raffinose and
stachyose combined of less than 14.5 ug/g, and a seed sucrose content
greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)

Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 84.2%; Score 1482; DB 2; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 54 ATGTCATCGAAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATT 113
| | | | |
Db 1 ATGTCATCGAAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATT 60
| | | | |
QY 114 CAGTCCGTGTACACTAGAAAACCGAATTTGTTCACGGAACAGGAATGGCACTAT 173
| | | | |
Db 61 CAGTCCGTGTACACTAGAAAACCGAATTTGTTCACGGAACAGGAATGGCACTAT 120
| | | | |
QY 174 CAGTGGATTGTCAAAACCCAAATCTGTCAAAATACGAATTTAAACCAATCCATTTCT 233
| | | | |
Db 121 CAGTGGATTGTCAAAACCCAAATCTGTCAAAATACGAATTTAAACCAATCCATTTCT 180
| | | | |
QY 234 AAATTAGGGGTAAATGCTTGTGGGTGGGTAAGCAACCGCTCAACCTTCAACCGGTGT 293
| | | | |
Db 181 AAATTAGGGGTAAATGCTTGTGGGTGGGTAAGCAACCGCTCAACCTTCAACCGGTGT 240
| | | | |
QY 294 GTTATTGCTAACCGAGAGGCAATTTATGAGGCTACAAAGACAGATTTCAACAGCCAA 353
| | | | |
Db 241 GTTATTGCTAACCGAGAGGCAATTTATGAGGCTACAAAGACAGATTTCAACAGCCAA 300
| | | | |
QY 354 TACTTGGGTCCCTCACCAGGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGAA 413
| | | | |
Db 301 TACTTGGGTCCCTCACCAGGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGAA 360
| | | | |
QY 414 ATCTATGCCCATTTCAAGAGCCGTGCTTCAATGTGTAACCTTGACGACATTTGTGGG 473
| | | | |
Db 361 ATCTATGCCCATTTCAAGAGCCGTGCTTCAATGTGTAACCTTGACGACATTTGTGGG 420
| | | | |
QY 474 GGATGGGATATCAGCAACATGAACTGTGCTGATGCTCATGGGCAAGGCAAGGTTTAC 533
| | | | |
Db 421 GGATGGGATATCAGCAACATGAACTGTGCTGATGCTCATGGGCAAGGCAAGGTTTAC 480
| | | | |
QY 534 ATCGATTTGCAAGACAGTTGAGGCTTCAATGGAATCCATGCTTCCACTCCCGGAATC 593
| | | | |
Db 481 ATCGATTTGCAAGACAGTTGAGGCTTCAATGGAATCCATGCTTCCACTCCCGGAATC 540
| | | | |
QY 594 TATGACCCGGATTTTATTGCTGCTGCAACCAAGAGAGCGTGCCCAACAGCTATCAAGG 653
| | | | |
Db 541 TATGACCCGGATTTTATTGCTGCTGCAACCAAGAGAGCGTGCCCAACAGCTATCAAGG 600
| | | | |
QY 654 ACAAGCAAGACAGTTCAACAAATCATCAAGACATCAAGACGTTAAAGAACCCACC 713
| | | | |
Db 601 ACAAGCAAGACAGTTCAACAAATCATCAAGACATCAAGACGTTAAAGAACCCACC 660
| | | | |
QY 714 AAAGTGCAAGAGGTTGATCTGTGACTGCTGCAACAGAGAGGTAGATATTTGGTT 773
| | | | |
Db 661 AAAGTGCAAGAGGTTGATCTGTGACTGCTGCAACAGAGAGGTAGATATTTGGTT 720
| | | | |
QY 774 GTGGGCTTTAATGACACCATGAGAGATCTTGTGCTGTGTGACAGAAATGAGGCTGAG 833
| | | | |
Db 721 GTGGGCTTTAATGACACCATGAGAGATCTTGTGCTGTGTGACAGAAATGAGGCTGAG 780
| | | | |
QY 834 ATTTTCTCTTCACTTGTATGCTTGTGTGTATGGAATATGTTCTTTCAATTAAT 893
| | | | |
Db 781 ATTTTCTCTTCACTTGTATGCTTGTGTGTATGGAATATGTTCTTTCAATTAAT 840
| | | | |
QY 894 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCATGCGAGAACT 953
| | | | |
Db 841 GGAAGCCCTCAGAACACTTTTGTATCAGAGGCTGATTTGATCTTGCATGCGAGAACT 900
| | | | |
QY 954 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTTGAT 1013
| | | | |
Db 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTTGAT 960
| | | | |
QY 1014 TTCTTTTGGGGGCTGTGATCAAGCAACATTTATAGTCAATTAAACCATTTGGGAAC 1073
| | | | |
Db 961 TTCTTTTGGGGGCTGTGATCAAGCAACATTTATAGTCAATTAAACCATTTGGGAAC 1020
| | | | |
QY 1074 AATGATGGATGATCTTTGGGCTCCAAACTTTCCGTTCCAAAGAAATCTCAAGAGC 1133
| | | | |
Db 1021 AATGATGGATGATCTTTGGGCTCCAAACTTTCCGTTCCAAAGAAATCTCAAGAGC 1080
| | | | |
QY 1134 AACGTTGTGATGATATGATCAACAGCAATGCCATCTCTATGAGCTGTGGAACATCCA 1193
| | | | |

AUTHORS Majee, M., Majumder, A.N.L. and Mundree, S.G.
TITLE Molecular characterization of Xvlnol, a myo-inositol-1-phosphate
JOURNAL synchase from the resurrection plant *Xerophyta vascosa*
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 1533)
TITLE Majee, M., Majumder, A.N.L. and Mundree, S.G.
JOURNAL Submitted (14-JUN-2003) Molecular and Cell Biology, University of
 Cape Town, Private Bag, Rondebosch, Western Cape 7701, South Africa
FEATURES Location/Qualifiers
 1. .1533
 /organism="Xerophyta vascosa"
 /mol_type="mRNA"
 /db_xref="taxon:90708"
 1. .1533
 /note="Xvlnol"
 /codon_start=1
 /product="myo-inositol-1-phosphate synthase INO1"
 /protein_id="AAP85531.1"
 /db_xref="GI:32492882"
 /translation="MTESFVRSFPHVRYGAEIESDYDTTELTVSHDASRWV
 RPKSVRYNFRRTTTPKLGVMVGWGNNGSTLTAGVIANREGISWATDKYQOANY
 GSIJTOASTIRVGSYNGEBEIVAPFKSLIPVNPBDLVFGMSIDMNLADAMTKAYLD
 IDLOKQLRPYMESMVPPLPGIYDPDPIAANGSRANVIGTKKEQKOEIIOIKREFKE
 KSKVDKVVYLVMTANTERSYNVGVGNDTMENTILASVDKRAEISPTVAIAICVMEGI
 PTNGSPQTFVPGILDLAIKKNCLIGGDFRSQGTAKMSVLVDPLVAGIKRPTSVS
 YNHLGNDGMSISAPQTFRSKEISKSNVDDVANGSLIYEGEPDHVVVIVKYVCV
 GDSKRAMDEYTSIEIFMGKNTIVMNTGDSLLAIPILIDVLVLAELSTRIOKAGE
 EKFSFHPVATLLSYLTKAPLVPGTPVNVNALAKORAMLENVIRACVGLAPENNMILE
 YK"

ORIGIN
 Query Match 1.5%; Score 26; DB 8; Length 1533;
 Best Local Similarity 100.0%; Pred. No. 0.024;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 AGCAATGCATCTCTATGAGCCTGG 1183
 1105 AGCAATGCATCTCTATGAGCCTGG 1130

RESULT 13
LOCUS AY096554 1564 bp mRNA linear PLN 18-SEP-2002
DEFINITION Arabidopsis thaliana putative myo-inositol-1-phosphate synthase
 (At5g10170) mRNA, complete cds.
ACCESSION AY096554
VERSION AY096554.1 GI:20465488
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1564)
 Yamada, K., Bahh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,
 Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
 Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,
 Carinci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
 Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
 Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,
 Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Arabidopsis Open Reading Frame (ORF) Clones
 2 (bases 1 to 1564)
 Yamada, K., Bahh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,
 Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
 Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,
 Carinci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
 Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
 Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,

TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN
 Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carinci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.
 The Salk, Stanford, PGEC (SSP) Consortium members constructed and
 sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada, K.,
 Bahh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,
 Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,
 Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,
 Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,
 Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P.,
 Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.
 Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
 to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP
 /PGEC) contributed equally to this work as PIs.
 Annotation is based on the January 2002 version of the Arabidopsis
 genome submitted to Genbank.
FEATURES Location/Qualifiers
 1. .1564
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="U018590"
 /ecotype="Columbia"
 /note="This clone is in PUNI 51."
 1. .1564
 /gene="At5g10170"
 1. .1533
 /gene="At5g10170"
 /codon_start=1
 /evidence=experimental
 /product="putative myo-inositol-1-phosphate synthase"
 /protein_id="AA020204.1"
 /db_xref="GI:20465489"
 /translation="MTESFVRSFPHVRYGAEIESDYDTTELTVSHDASRWV
 RPKSVRYNFRRTTTPKLGVMVGWGNNGSTLTAGVIANREGISWATDKYQOANY
 GSIJTOASTIRVGSYNGEBEIVAPFKSLIPVNPBDLVFGMSIDMNLADAMTKAYLD
 IDLOKQLRPYMESMVPPLPGIYDPDPIAANGSRANVIGTKKEQKOEIIOIKREFKE
 KSKVDKVVYLVMTANTERSYNVGVGNDTMENTILASVDKRAEISPTVAIAICVMEGI
 PTNGSPQTFVPGILDLAIKKNCLIGGDFRSQGTAKMSVLVDPLVAGIKRPTSVS
 YNHLGNDGMSISAPQTFRSKEISKSNVDDVANGSLIYEGEPDHVVVIVKYVCV
 GDSKRAMDEYTSIEIFMGKNTIVMNTGDSLLAIPILIDVLVLAELSTRIOKAGE
 EKFSFHPVATLLSYLTKAPLVPGTPVNVNALAKORAMLENVIRACVGLAPENNMILE
 YK"
 1534. .1564
 /gene="At5g10170"

ORIGIN
 3' UTR

QY 966 GATGACTTCAGAGTGTGACACCA 991
 913 GATGACTTCAGAGTGTGACACCA 938

RESULT 14
LOCUS AX054630 1781 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 1 from Patent WO0073473.
ACCESSION AX054630
VERSION AX054630.1 GI:12228190
KEYWORDS

QY 966 GATGACTTCAGAGTGTGACACCA 991
 913 GATGACTTCAGAGTGTGACACCA 938

Query Match 1.5%; Score 26; DB 8; Length 1564;
 Best Local Similarity 100.0%; Pred. No. 0.024;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1
AUTHORS Georges,F., Hussain,A.A. and Keller,W.A.
TITLE Method for reducing phytate in canola meal using genetic
manipulation involving myo-inositol 1-phosphate synthase gene
JOURNAL Patent: WO 0073473-A 1 07-DEC-2000;
NATIONAL RESEARCH COUNCIL CANADA (CA)
FEATURES
source Location/Qualifiers
1..1781
/organism="Brassica napus"
/mol_type="unassigned DNA"
/db_xref="taxon:3708"
ORIGIN
Query Match 1.5%; Score 26; DB 6; Length 1781;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 960 GGTGAGATGACTTCAAGAGTGCTCA 985
|||||
DB 963 GGTGAGATGACTTCAAGAGTGCTCA 988
RESULT 15
BN066307 1781 bp mRNA linear PLN 02-MAR-2000
LOCUS Brassica napus myo-inositol 1-phosphate synthase mRNA, complete
DEFINITION cds.
ACCESSION U66307
VERSION U66307.1 GI:1513227
KEYWORDS
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 1781)
AUTHORS Hussain,A., Bourgeois,J., Polvi,S., Tsang,E., Keller,W.A. and
Georges,F.
TITLE Cloning of a full length cDNA encoding myo-inositol 1-phosphate
synthase from Brassica napus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1781)
AUTHORS Hussain,A., Bourgeois,J., Polvi,S., Tsang,E., Keller,W.A. and
Georges,F.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1996) National Research Council of Canada, Plant
Biotechnology Institute, 110 Gymnasium Place, Saskatoon,
Saskatchewan S7N 0W9, Canada
FEATURES
source Location/Qualifiers
1..1781
/organism="Brassica napus"
/mol_type="mRNA"
/db_xref="taxon:3708"
57..1589
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAB06756.2"
/db_xref="GI:7144546"
/translation="MFIESPKYSEPNVKTENSHSVYDETTEVYHENVNGAYQWIV
KRVVYDFKTDIVRPKLVGLVWGNGNGSTLTAGVIANKEGISWATKQVQANVF
GSLTQASIRVSEFNGEEMVAPFKSLVPMVNPDDVVGMDISDMVLADAMGRARVL
IDLOKQLRPYMENIVLPGLIYDPDFIAAQGRANNVIGTKKEQVDOIIKDMREFKS
KKKQDVVVLMTANTERYGNVIGLNDTMEINMNSVDRDESEISPTLVAIACVLEGI
PRINGSRPTFVPGILDLAIKNNVLIGDDDFKSCOTKMSVLVDPLVAGAGIKPTGISV
YVHLGNNDGMNLSAPOTFRSKLSKSNVVDNVAASGILLFEGEHPDHVYVVKYPIV
AUSKRAMDERTSEIFWGNKTVIMHNTCEDSLAAPIILDVLVAISTRIQFSKEKE
GKFHSPVATKLSYLTKAPLVPGTVPVNALSKORAMLENTLRACVGLAPENNMLE
YK"

ORIGIN
Query Match 1.5%; Score 26; DB 8; Length 1781;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 960 GGTGAGATGACTTCAAGAGTGCTCA 985
|||||
DB 963 GGTGAGATGACTTCAAGAGTGCTCA 988
Search completed: June 8, 2005, 11:26:37
Job time : 7685.29 secs


```
Db 1081 AACGTTGATGATATGTCACACAGCAATGCATCTCTATAGCTGGTGAACTCCA 1140
Qy 1194 GACCATGTTGTTTATTAAATGATGTCCTTACGTAGGGGACAGCAAGAGCCATGAT 1253
Db 1141 GACCATGTTGTTTATTAAATGATGTCCTTACGTAGGGGACAGCAATAGAGCCATGAT 1200
Qy 1254 GAGTACACTTCAGAGATATTCATGAGGTGGAAGAGACCACTGTTTGCACACATGCT 1313
Db 1201 GAGTACACTTCAGAGATATTCATGAGGTGGAAGAGACCACTGTTTGCACACATGCT 1260
Qy 1314 GAGGATTCCTCTTACCTGCTCTTATTTATCTTGAGCTGGTCTTCTGTCAGTCAAG 1373
Db 1261 GAGGATTCCTCTTACCTGCTCTTATTTATCTTGAGCTGGTCTTCTGTCAGTCAAG 1320
Qy 1374 ACTAGATGAGTTTAAAGCTGAAATGAGGAAAAATTCACATCTATTCACCAAGTTGCT 1433
Db 1321 ACTAGATGAGTTTAAAGCTGAAATGAGGAAAAATTCACATCTATTCACCAAGTTGCT 1380
Qy 1434 ACCATCTGAGTACCTCAACCAAGGCTCTCTGTTCCACCGGTTACACAGTGTGAT 1493
Db 1381 ACCATCTGAGTACCTCAACCAAGGCTCTCTGTTCCACCGGTTACACAGTGTGAT 1440
Qy 1494 GCATTGTCAAAGCAGCGTGCATGCTGGAATAATGAGGCTTGTGTTGATTGGCC 1553
Db 1441 GCATTGTCAAAGCAGCGTGCATGCTGGAATAATGAGGCTTGTGTTGATTGGCC 1500
Qy 1554 CCAGAGATPAACATGATTTCTCGAGTCAAGTGA 1586
Db 1501 CCAGAGATPAACATGATTTCTCGAGTCAAGTGA 1533
```

RESULT 7

ADQ14494

ADQ14494 standard; cDNA; 1533 BP.

XX ADQ14494;

XX 23-SEP-2004 (first entry)

XX Mutant soybean myo-inositol 1-phosphate synthase cDNA #1.

XX Soybean: myo-inositol 1-phosphate synthase; gene; ss;

XX myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;

XX inorganic phosphate; mutant.

XX Glycine max.

XX Synthetic.

XX Key Location/Qualifiers

XX CDS 1..1533

XX /product= "Mutant soybean myo-inositol 1-phosphate

XX mutation replace(1241,G)

XX /tag= b

XX US2004128713-A1.

XX 01-JUL-2004.

XX 21-NOV-2003; 2003US-00718952.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX 26-APR-1999; 99US-00293315.

XX 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX (GRACE/) GRACE D J.

XX (STREIT/) STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX MPI; 2004-533135/51.
DR P-PSDB; ADQ14495.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX Claim 10; SEQ ID NO 5; 48bp; English.
PS
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 84.2%; Score 1482; DB 12; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 54 ATGTTTCATGAGATTTTAAAGTTGATGCTCTTATGTAAGTACACCGAGCTGAGAT 113
Db 1 ATGTTTCATGAGATTTTAAAGTTGATGCTCTTATGTAAGTACACCGAGCTGAGAT 60
Qy 114 CAGTCCGTTACACCTACGAAACCAACCGAATCTGTTACAGAGACGAAATGACCTAT 173
Db 61 CAGTCCGTTACACCTACGAAACCAACCGAATCTGTTACAGAGACGAAATGACCTAT 120
Qy 174 CAGTGGATTGCAAAACCAATCTGTCAATGCAATTTAAACCAACATCATGTTCT 233
Db 121 CAGTGGATTGCAAAACCAATCTGTCAATGCAATTTAAACCAACATCATGTTCT 180
Qy 234 AAATTAGGGGTAAATGCTGTTGGGTGGGTGGAACCAACGCTCAACCGTGT 293
Db 181 AAATTAGGGGTAAATGCTGTTGGGTGGGTGGAACCAACGCTCAACCGTGT 240
Qy 294 GTTATTGCTAACCGAGAGGCAATTATGAGGCTTACAAAGAGCAAGATTCAGCAAT 353
Db 241 GTTATTGCTAACCGAGAGGCAATTATGAGGCTTACAAAGAGCAAGATTCAGCAAT 300
Qy 354 TACTTTGGCTCCCTACCAACCAAGCTCAGCTATCCAGTTGGTCTTCCAGGAGAGAA 413
Db 301 TACTTTGGCTCCCTACCAACCAAGCTCAGCTATCCAGTTGGTCTTCCAGGAGAGAA 360
Qy 414 ATCTATGCCCACTTAAAGAGCTGCTTCAATGTTTAAACCTGACGACATTTGTTGGG 473
Db 361 ATCTATGCCCACTTAAAGAGCTGCTTCAATGTTTAAACCTGACGACATTTGTTGGG 420
Qy 474 GGATGGAGATATGAGCAATGAACTGGCTGATGCTGATGCGCAGGAGCAAGGATTTGAC 533
Db 421 GGATGGAGATATGAGCAATGAACTGGCTGATGCTGATGCGCAGGAGCAAGGATTTGAC 480
Qy 534 ATGATTTTGCAAGAACAGTTGAGGCTTATCATGAAATCATGCTTCACTCCCGGAATC 593
Db 481 ATGATTTTGCAAGAACAGTTGAGGCTTATCATGAAATCATGCTTCACTCCCGGAATC 540
```

```

QY 594 TATGACCCGGATTTTCATTGCTGCCAACGAAGAGAGCGTGCCAAACACGTCATCAAGGCG 653
|
|
|
Db 541 TATGACCCGGATTTTCATTGCTGCCAACGAAGAGAGCGTGCCAAACACGTCATCAAGGCG 600
|
|
|
QY 654 ACMAAGCAAGACCAAGTTCACAAATCATCAAAAGACATCAAGCGCTTTAAGNAGCCACC 713
|
|
|
Db 601 ACAAAAGCAAGACCAAGTTCACAAATCATCAAAAGACATCAAGCGCTTTAAGNAGCCACC 660
|
|
|
QY 714 AAAGTGAACAAGGTGTTTACCTGTCGCAACACAGAGAGTACAGTAATTTGGTT 773
|
|
|
Db 661 AAAGTGAACAAGGTGTTTACCTGTCGCAACACAGAGAGTACAGTAATTTGGTT 720
|
|
|
QY 774 GTGGGCTTTAATGACCAATGGAAGAACTCTTGGCTGCTGTGACAGAAATAGCGCTGAG 833
|
|
|
Db 721 GTGGGCTTTAATGACCAATGGAAGAACTCTTGGCTGCTGTGACAGAAATAGCGCTGAG 780
|
|
|
QY 834 ATTTCTCTTCCACCTTGATGCCATGCTGTTGTTAAGGAAATGTTCTTTCATTTAT 893
|
|
|
Db 781 ATTTCTCTTCCACCTTGATGCCATGCTGTTGTTAAGGAAATGTTCTTTCATTTAT 840
|
|
|
QY 894 GGAAGCCCTCAGAACACTTTTGTACCAAGGGCTGATGATCTTGCCATCGCGAGAACT 953
|
|
|
Db 841 GGAAGCCCTCAGAACACTTTTGTACCAAGGGCTGATGATCTTGCCATCGCGAGAACT 900
|
|
|
QY 954 TTGATTGTGAGAGTACTTCAAGAGTGTCAAGCCAAATGAAATCTGTGTGTTGAT 1013
|
|
|
Db 901 TTGATTGTGAGAGTACTTCAAGAGTGTCAAGCCAAATGAAATCTGTGTGTTGAT 960
|
|
|
QY 1014 TTCTTGTGAGGGCTGTGATCAAGCCAAATCTATAGTCAAGTTCACAAATCTGGGAAAC 1073
|
|
|
Db 961 TTCTTGTGAGGGCTGTGATCAAGCCAAATCTATAGTCAAGTTCACAAATCTGGGAAAC 1020
|
|
|
QY 1074 AATGATGTATGATATCTTTCGCGTCCACAAACTTTCCGTTCCAAAGGAAATCTCAAGAGC 1133
|
|
|
Db 1021 AATGATGTATGATATCTTTCGCGTCCACAAACTTTCCGTTCCAAAGGAAATCTCAAGAGC 1080
|
|
|
QY 1134 AAGTGTGTATGATATGTGTCAACAGCAATGCCATCTCTATAGAGCTGTGTGAATCCA 1193
|
|
|
Db 1081 AAGTGTGTATGATATGTGTCAACAGCAATGCCATCTCTATAGAGCTGTGTGAATCCA 1140
|
|
|
QY 1194 GACCATGTTGTTTATTAAGTATGTGCTTAAAGTGGGACAGCAAGAGCCATGAT 1253
|
|
|
Db 1141 GACCATGTTGTTTATTAAGTATGTGCTTAAAGTGGGACAGCAATGAGCCATGAT 1200
|
|
|
QY 1254 GAGTACACTTCAGAGATTTTCAATGSGTGAAGAGACCACTGTTTTCACAAACATGTC 1313
|
|
|
Db 1201 GAGTACACTTCAGAGATTTTCAATGSGTGAAGAGACCACTGTTTTCACAAACATGTC 1260
|
|
|
QY 1314 GAGGATTTCCCTTCTTACCTGCTCTATTTATCTTGAGCTGTGTCTTGTGAGCTCAGC 1373
|
|
|
Db 1261 GAGGATTTCCCTTCTTACCTGCTCTATTTATCTTGAGCTGTGTCTTGTGAGCTCAGC 1320
|
|
|
QY 1374 ACTAGAAATGAGATTTAAAGTGAAGTGAAGGAAATTCACATTCATCCACCAATGTT 1433
|
|
|
Db 1321 ACTAGAAATGAGATTTAAAGTGAAGTGAAGGAAATTCACATTCATCCACCAATGTT 1380
|
|
|
QY 1434 ACCATCTCAGTACCTCAGCAAGAGCTCTCTGTGTTCCACCGGGTACACAGTGTGAAT 1493
|
|
|
Db 1381 ACCATCTCAGTACCTCAGCAAGAGCTCTCTGTGTTCCACCGGGTACACAGTGTGAAT 1440
|
|
|
QY 1494 GCATTGTCAAAAGCAGGTGCAATGCTGGAATAATGAGGCGTTGTGTGATTGGCC 1553
|
|
|
Db 1441 GCATTGTCAAAAGCAGGTGCAATGCTGGAATAATGAGGCGTTGTGTGATTGGCC 1500
|
|
|
QY 1554 CCAGAGAAATTAACATGATTTCTCGAGTCAAGTGA 1586
|
|
|
Db 1501 CCAGAGAAATTAACATGATTTCTCGAGTCAAGTGA 1533
|
|
|

```

RESULT 8
ADS81997
ID ADS81997 standard; cDNA; 1533 BP.
XX

```

AC ADS81997;
XX
XX 18-NOV-2004 (first entry)
DT
XX
XX Soybean myo-inositol 1-phosphate synthase cDNA mutant #1.
DE
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phylic acid;
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
XX
XX Glycine max; line LR33.
OS
XX Synthetic.
FH
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
FT mutation replace(1188,G)
FT /*tag= b
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
PI
XX WPI: 2004-639957/62.
XX DR P-PsDB; ADS81998.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphoric synthase, useful for altering raffinose saccharide, sucrose,
PT phylic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 8; SEQ ID NO 5; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphoric synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phylic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g), provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phylic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a mutant myo-
XX inositol 1-phosphate synthase.
XX

```

Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 84.2%; Score 1482; DB 13; Length 1533;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 54 ATGTTCATCGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCCGAGCTGAGATT 113
DB 1 ATGTTCATCGAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCCGAGCTGAGATT 60
QY 114 CAGTCCGTGTACAACTACGAAACACCGAATTGTTCAAGAGAACAGAAATGACCTAT 173
DB 61 CAGTCCGTGTACAACTACGAAACACCGAATTGTTCAAGAGAACAGAAATGACCTAT 120
QY 174 CAGTGTATGTCAAACCCAAATCTGTCAATACGAATTTAAACCAACATCATGTTCTT 233
DB 121 CAGTGTATGTCAAACCCAAATCTGTCAATACGAATTTAAACCAACATCATGTTCTT 180
QY 234 AATTTAGGGGTATGTGTGGGTGGGTGGGAAACACCGCTCAACCCCTCACGGTGT 293
DB 181 AATTTAGGGGTATGTGTGGGTGGGTGGGAAACACCGCTCAACCCCTCACGGTGT 240
QY 294 GTTATGTCTAACCGAGAGGCTATTCATGCGCTACAAAGGACAAAGATTCAACAGCAAT 353
DB 241 GTTATGTCTAACCGAGAGGCTATTCATGCGCTACAAAGGACAAAGATTCAACAGCAAT 300
QY 354 TACTTTGGCTCCCTCACCCAGGCTCACTATCCGAGTTGGTCTTCCAGGAGAGAA 413
DB 301 TACTTTGGCTCCCTCACCCAGGCTCACTATCCGAGTTGGTCTTCCAGGAGAGAA 360
QY 414 ATCTATGCCCATTCAGAGGCTGTCTTCCAAATGTTAAACCTGACGACATGTTGTTGG 473
DB 361 ATCTATGCCCATTCAGAGGCTGTCTTCCAAATGTTAAACCTGACGACATGTTGTTGG 420
QY 474 GGTGAGATATACGAAACATGAACCTGTGATGCATGGCCAGGCAAGAGTGTGAC 533
DB 421 GGTGAGATATACGAAACATGAACCTGTGATGCATGGCCAGGCAAGAGTGTGAC 480
QY 534 ATCGATTTCAGAGAGATTGAGGCTTTCATGGAATTCATGCTTCCACTCCCGGAATC 593
DB 481 ATCGATTTCAGAGAGATTGAGGCTTTCATGGAATTCATGCTTCCACTCCCGGAATC 540
QY 594 TATGACCCGGATTTCATGCTGTGCAACGAGAGACGCTGCCAAACGTCATCAAGGCG 653
DB 541 TATGACCCGGATTTCATGCTGTGCAACGAGAGACGCTGCCAAACGTCATCAAGGCG 600
QY 654 ACAAGCAGAGAGATTCAACAATCATCAAGACATCAAGGCTTTAAGAGAGCAC 713
DB 601 ACAAGCAGAGAGATTCAACAATCATCAAGACATCAAGGCTTTAAGAGAGCAC 660
QY 714 AAAGTGAACAAGTGTGTGACTGTGACTGCCAACAAGAGATACAGTAATTTGGTT 773
DB 661 AAAGTGAACAAGTGTGTGACTGTGACTGCCAACAAGAGATACAGTAATTTGGTT 720
QY 774 GTGGGCTTAAAGACCAATGAGAAATCTTGGCTGTGACGAGAAATGAGGCTGAG 833
DB 721 GTGGGCTTAAAGACCAATGAGAAATCTTGGCTGTGACGAGAAATGAGGCTGAG 780
QY 834 ATTTCTCTTCCACCTTGTATGCTGTTGTGATGAAATTTCTCTTCAATTAAT 893
DB 781 ATTTCTCTTCCACCTTGTATGCTGTTGTGATGAAATTTCTCTTCAATTAAT 840
QY 894 GGAAGCCTTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTGCATGCGAGAACCT 953
DB 841 GGAAGCCTTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTGCATGCGAGAACCT 900
QY 954 TTGATTTGTGAGATGACTTCAAGAGTGTGAGAACCAAAATGAATCTGTGTTGTTAT 1013
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGAGAACCAAAATGAATCTGTGTTGTTAT 960
QY 1014 TTCTCTTGTGGGGCTGTGTATCAAGCCACATCTATAGTCAACCATCTGGGAAC 1073
DB 961 TTCTCTTGTGGGGCTGTGTATCAAGCCACATCTATAGTCAACCATCTGGGAAC 1020

```

```

QY 1074 AATGATGATATGATCTTTTCGGCTCCACAAACTTTCCGTTCCAGGAATCTCCAGAGC 1133
DB 1021 AATGATGATATGATCTTTTCGGCTCCACAAACTTTCCGTTCCAGGAATCTCCAGAGC 1080
QY 1134 AACGTTGTGATATGATGTCACACGACATGCTCTATAGGCTGGTGAACATCCA 1193
DB 1081 AACGTTGTGATATGATGTCACACGACATGCTCTATAGGCTGGTGAACATCCA 1140
QY 1194 GACCATGTTGTTTATTAATGATGTGCTTACGTAGGGGACAGCAAGAGCCATGAT 1253
DB 1141 GACCATGTTGTTTATTAATGATGTGCTTACGTAGGGGACAGCAATGAGCCATGAT 1200
QY 1254 GAGTACACTTCAGAGATATTCATGSGTGAAAGAGACCAATGTTTTCACACATGC 1313
DB 1201 GAGTACACTTCAGAGATATTCATGSGTGAAAGAGACCAATGTTTTCACACATGC 1260
QY 1314 GAGGATTCCTCTTACTGCTCTATATCTTGAGACTTGTGCTTGTGAGCTCAGC 1373
DB 1261 GAGGATTCCTCTTACTGCTCTATATCTTGAGACTTGTGAGCTTGTGAGCTCAGC 1320
QY 1374 ACTAGATGAGATTTAAAGCTGAAATGAGGAAATTCACATTCACACCATGTTGCT 1433
DB 1321 ACTAGATGAGATTTAAAGCTGAAATGAGGAAATTCACATTCACACCATGTTGCT 1380
QY 1434 ACCATCTCAGTACTCTCAACAGGCTCTGTGTTCCACCGGATACCAAGTGTGAAT 1493
DB 1381 ACCATCTCAGTACTCTCAACAGGCTCTGTGTTCCACCGGATACCAAGTGTGAAT 1440
QY 1494 GCATTGTCAAAGCAGGTCATAGCTGSAATCATTAAGGGCTGTGTGGAATTGGGC 1553
DB 1441 GCATTGTCAAAGCAGGTCATAGCTGSAATCATTAAGGGCTGTGTGGAATTGGGC 1500
QY 1554 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1586
DB 1501 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1533

```

RESULT 9
 ADQ14504
 ID ADQ14504 standard; cDNA; 1533 BP.
 XX
 AC ADQ14504;
 XX
 DT 23-SEP-2004 (first entry)
 XX
 DE Wild type soybean myo-inositol 1-phosphate synthase cDNA #2.
 XX
 KW Soybean; myo-inositol 1-phosphate synthase; gene; ss;
 KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
 KW inorganic phosphate.
 XX
 OS Glycine max.
 XX
 PH Key Location/Qualifiers
 FT CDS 1..1533
 FT /tag= a
 FT /product= "wild type soybean myo-inositol 1-phosphate
 FT synthase #2"
 FT
 PN US2004128713-A1.
 XX
 XX
 PD 01-JUL-2004.
 XX
 XX 21-NOV-2003; 2003US-00718952.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98MO-US006822.
 PR 26-APR-1999; 99US-0029315.
 PR 11-MAR-2002; 2002US-00025003.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
DR WPI; 2004-533135/51.
DR P-PSDB; ADQ14505.
XX
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX
PS Claim 4; SEQ ID NO 15; 48bp; English.
XX
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 8.3%; Score 146; DB 12; Length 1533;
Best Local Similarity 100.0%; Pred. No. 2.3e-60;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 870 ATGAAAATGTTCTTCTTATTAATGAAGCCCTCAGAACATTTTGTACAGGGCTGATT 929
DB 817 ATGAAAATGTTCTTCTTATTAATGAAGCCCTCAGAACATTTTGTACAGGGCTGATT 876
QY 930 GATCTTGCATCGCGAGAACACTTTGATTGGTGAGATGACTTCAAGAGTGTACAGCC 989
DB 877 GATCTTGCATCGCGAGAACACTTTGATTGGTGAGATGACTTCAAGAGTGTACAGCC 936
QY 990 AAAATGAATCTGTGTGTGTTGATT 1015
DB 937 AAAATGAATCTGTGTGTGTTGATT 962

RESULT 10
ADQ14500
ID ADQ14500 standard; cDNA; 1533 BP.
XX
XX ADQ14500;
XX
XX 23-SEP-2004 (first entry)
XX
XX Mutant soybean myo-inositol 1-phosphate synthase cDNA #3.
DE
XX
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KM inorganic phosphate; mutant.
XX
XX Glycine max.
OS Synthetic.
OS
FH Key Location/Qualifiers
FT CDS 1..1533
FT /*tag= a

FT
FT
XX
XX
PN /product= "Mutant soybean myo-inositol 1-phosphate
XX synthase #3"
XX
XX US2004128713-A1.
XX
XX
PD 01-JUL-2004.
XX
XX
PF 21-NOV-2003; 2003US-00718952.
XX
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006882.
PR 26-APR-1999; 99US-00289315.
PR 11-MAR-2002; 2002US-00025003.
XX
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
DR WPI; 2004-533135/51.
DR P-PSDB; ADQ14501.
XX
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX
PS Claim 10; SEQ ID NO 11; 48bp; English.
XX
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

Query Match 8.3%; Score 146; DB 12; Length 1533;
Best Local Similarity 100.0%; Pred. No. 2.3e-60;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 870 ATGAAAATGTTCTTCTTATTAATGAAGCCCTCAGAACATTTTGTACAGGGCTGATT 929
DB 817 ATGAAAATGTTCTTCTTATTAATGAAGCCCTCAGAACATTTTGTACAGGGCTGATT 876
QY 930 GATCTTGCATCGCGAGAACACTTTGATTGGTGAGATGACTTCAAGAGTGTACAGCC 989
DB 877 GATCTTGCATCGCGAGAACACTTTGATTGGTGAGATGACTTCAAGAGTGTACAGCC 936
QY 990 AAAATGAATCTGTGTGTGTTGATT 1015
DB 937 AAAATGAATCTGTGTGTGTTGATT 962

RESULT 11
ADQ14502
ID ADQ14502 standard; cDNA; 1533 BP.
XX

AC ADQ14502;
XX
XX 23-SEP-2004 (first entry)
XX
XX Mutant soybean myo-inositol 1-phosphate synthase cDNA #4.
DE
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate; mutant.
XX
XX Glycine max.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #4"
XX
XX US2004128713-A1.
XX
XX 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
XX PF 07-APR-1998; 98WO-US006822.
XX PR 26-APR-1999; 99US-0029315.
XX 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX (GRAC/) GRACE D J.
XX (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX WPI; 2004-53135/51.
XX P-PSDB; ADQ14503.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX Example 8; SEQ ID NO 13; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents cDNA encoding a mutant soybean
XX myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
XX
XX Query Match 8.3%; Score 146; DB 12; Length 1533;
XX Best Local Similarity 100.0%; Pred. No. 2.3e-60;
XX Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 870 ATGGAATATGTTCTTCAATATGAGACCCCTCAGAACACTTTGTACGAGGCTGATT 929

Db 817 ATGGAATATGTTCTTCAATATGAGACCCCTCAGAACACTTTGTACGAGGCTGATT 876
Qy 930 GATCTTGCATGCGGAGAACACTTTGATTGTGAGATGATCTTCAAGAGTGTACAGCC 989
Db 877 GATCTTGCATGCGGAGAACACTTTGATTGTGAGATGATCTTCAAGAGTGTACAGCC 936
Qy 990 AAAATGAATCTGTGTTGTTGATT 1015
Db 937 AAAATGAATCTGTGTTGTTGATT 962
RESULT 12
ID ADS82005 standard; cDNA; 1533 BP.
XX
XX ADS82005;
XX
XX 18-NOV-2004 (first entry)
XX
XX Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 2.
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
XX Glycine max; cultivar Wye.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
XX P-PSDB; ADS82006.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 2; SEQ ID NO 15; 34bp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
XX soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX phosphate synthase having decreasing capacity for the synthesis for myo-
XX inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX complement, subfragment or the complement of the subfragment, operably
XX linked to suitable regulatory sequences, where expression of the chimeric
XX gene results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX comprising the chimeric gene (with a heritable phenotype of a seed phytic
XX acid content of less than 17 micromol/g, a seed content of raffinose plus
XX stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX greater than 200 micromol/g, provided that the plant is not LR33), seeds
XX from the plant, making a soybean plant with the heritable phenotype
XX (comprising crossing LR33 or the plant comprising the chimeric gene with
XX an elite soybean plant and selecting a progeny plant of the cross of
XX crossing step that has a heritable phenotype as mentioned above), seeds
XX of soybean plant made by the above method, a soy protein product derived

CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with a
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytyc acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.

SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 8.3%; Score 146; DB 13; Length 1533;

Best Local Similarity 100.0%; Pred. No. 2.3e-60; Mismatches 0; Gaps 0;

Matches 146; Conservative 0; Indels 0; Gaps 0;

QY 870 ATGAAATGTTCTTTCATTATGAGACCCCTCAGAACCTTTGTACAGGGCTGATT 929

DB 817 ATGAAATGTTCTTTCATTATGAGACCCCTCAGAACCTTTGTACAGGGCTGATT 876

QY 930 GATCTTGCATGCGAGGAACAACCTTGTATGATGATGACTTCAAGATGCTCAGACC 989

DB 877 GATCTTGCATGCGAGGAACAACCTTGTATGATGATGACTTCAAGATGCTCAGACC 936

QY 990 AAAATGAATCTGTGTGTTGTTGATT 1015

DB 937 AAAATGAATCTGTGTGTTGTTGATT 962

RESULT 13

ADS82003 standard; cDNA; 1533 BP.

AC ADS82003;

DT 18-NOV-2004 (first entry)

XX Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 4.

XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytyc acid;

KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.

OS Glycine max; line 29018JF03.

XX Key Location/Qualifiers

FT CDS 1..1533

FT /tag= a

FT /product= "myo-inositol 1-phosphate synthase"

XX US2003074685-A1.

XX 11-MAR-2002; 2002US-00025003.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX Hitz WD, Sebastian SA;

XX WPI; 2004-63957/62.

XX P-PSDB; ADS82004.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-

PT phosphate synthase, useful for altering raffinose saccharide, sucrose,

XX phytyc acid and inorganic phosphate content of soybean seeds.

PS Example 8; SEQ ID NO 13; 34pp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis of myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment), operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytyc
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with a
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytyc acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.

SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 8.3%; Score 146; DB 13; Length 1533;

Best Local Similarity 100.0%; Pred. No. 2.3e-60; Mismatches 0; Gaps 0;

Matches 146; Conservative 0; Indels 0; Gaps 0;

QY 870 ATGAAATGTTCTTTCATTATGAGACCCCTCAGAACCTTTGTACAGGGCTGATT 929

DB 817 ATGAAATGTTCTTTCATTATGAGACCCCTCAGAACCTTTGTACAGGGCTGATT 876

QY 930 GATCTTGCATGCGAGGAACAACCTTGTATGATGATGACTTCAAGATGCTCAGACC 989

DB 877 GATCTTGCATGCGAGGAACAACCTTGTATGATGATGACTTCAAGATGCTCAGACC 936

QY 990 AAAATGAATCTGTGTGTTGTTGATT 1015

DB 937 AAAATGAATCTGTGTGTTGTTGATT 962

RESULT 14

ADS82001 standard; cDNA; 1533 BP.

AC ADS82001;

DT 18-NOV-2004 (first entry)

XX Soybean myo-inositol 1-phosphate synthase cDNA mutant #2.

XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytyc acid;

KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.

XX Glycine max; line 29010CP01.

XX Synthetic.

XX Key Location/Qualifiers

FT CDS 1..1533

FT /tag= a

FT /product= "myo-inositol 1-phosphate synthase"

FT mutation replace(260,G)

```

FT      /*tag= b
PN      US2003074685-A1.
XX
XX      17-APR-2003.
PD
XX      11-MAR-2002; 2002US-00025003.
PF
XX      08-APR-1997; 97US-00835751.
PR      07-APR-1998; 98WO-US006822.
XX
XX      (HITZ/) HITZ W D.
PA      (SEBA/) SEBASTIAN S A.
XX
XX      Hitz WD, Sebastian SA;
PI
XX      MPI; 2004-639957/62.
XX      P-PSDB; ADS82002.
XX
XX      Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT      phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT      phytic acid and inorganic phosphate content of soybean seeds.
XX
XX      Claim 8; SEQ ID NO 11; 34pp; English.
PS
XX
CC      The invention relates to an isolated nucleic acid fragment encoding a
CC      soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC      phosphate synthase having decreasing capacity for the synthesis for myo-
CC      inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC      nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC      complement, subfragment or the complement of the subfragment, operably
CC      linked to suitable regulatory sequences, where expression of the chimeric
CC      gene results in a decrease in expression of an endogenous or native gene
CC      encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC      comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC      acid content of less than 17 micromol/g, a seed content of raffinose plus
CC      stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC      greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC      from the plant, making a soybean plant with the heritable phenotype
CC      (comprising crossing LR33 or the plant comprising the chimeric gene with
CC      an elite soybean plant and selecting a progeny plant of the cross of
CC      crossing step that has a heritable phenotype as mentioned above), seeds
CC      of soybean plant made by the above method, a soy protein product derived
CC      from seeds of a soybean plant (homozygous for one or more gene encoding a
CC      mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC      the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC      heritable phenotype as mentioned above), and making or producing a
CC      soybean protein product derived from seeds of a soybean plant with
CC      heritable phenotype as mentioned above. The nucleic acid is useful for
CC      altering raffinose saccharide, sucrose, phytic acid and inorganic
CC      phosphate content of soybean seeds thus leading to valuable and useful
CC      soybean products, since the presence of high concentration of raffinose
CC      oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC      when consumed by humans. The present sequence encodes a mutant myo-
CC      inositol 1-phosphate synthase.
XX
XX      Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;
SQ
XX
Query Match      8.3%; Score 146; DB 13; Length 1533;
Best Local Similarity 100.0%; Pred. No. 2.3e-60;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      870 ATGGAATATGTTCTTTCATTAAATGGAAGCCTCAGAAACATTTTGACAGGCTGATT 929
XX      |||||||
XX      DB      817 ATGGAATATGTTCTTTCATTAAATGGAAGCCTCAGAAACATTTTGACAGGCTGATT 876
XX
XX      QY      930 GATCTGGCATGCGAGAGAACTTGATTTGGTGAGAGTACTTCAAGAGTGTTCAGACC 989
XX      |||||||
XX      DB      877 GATCTGGCATGCGAGAGAACTTGATTTGGTGAGAGTACTTCAAGAGTGTTCAGACC 936
XX
XX      QY      990 AAAATGAATCTGTGTGTTGATT 1015
XX      |||||||
XX      DB      937 AAAATGAATCTGTGTGTTGATT 962

```

```

RESULT 15
ADP93370
ID      ADP93370 standard; cDNA; 377 BP.
XX
XX      AC      ADP93370;
XX
XX      DT      09-SEP-2004 (first entry)
XX
XX      XX      Cotton expressed sequence tag, EST, #2381.
XX
XX      DE      Cotton; ss; EST; expressed sequence tag; plant; plant protection;
XX      KM      plant improvement; marker-assisted breeding.
XX
XX      OS      Gossypium hirsutum; variety Nucleon33B.
XX
XX      PN      US2004123338-A1.
XX
XX      PD      24-JUN-2004.
XX
XX      PF      08-DEC-2000; 2000US-00732627.
XX
XX      PR      10-DEC-1999; 99US-0170255P.
XX
XX      PA      (FINC/) FINCHER K L.
XX      PI      Fincher KL;
XX
XX      DR      MPI; 2004-479807/45.
XX
XX      PT      New substantially purified nucleic acid molecule that encodes a cotton
XX      protein or its fragment, useful as molecular tool for the targeting and
XX      isolation of novel genes for plant protection and improvement.
XX
XX      PS      Claim 1; SEQ ID NO 2381; 30pp; English.
XX
XX      CC      The invention relates to a substantially purified nucleic acid molecule
XX      CC      that encodes a cotton protein or its fragment comprising an EST
XX      CC      (expressed sequence tag) appearing as ADP90990-ADP95919. Also included
XX      CC      are a substantially purified cotton protein or its fragment encoded by a
XX      CC      nucleic acid molecule above and a transformed plant (having a nucleic
XX      CC      acid molecule which comprises: an exogenous promoter region which
XX      CC      functions in a plant cell to cause the production of a mRNA molecule; a
XX      CC      structural nucleic acid molecule comprising one of the ESTs or their
XX      CC      complements; a 3' non-translated sequence that functions in the plant
XX      CC      cell to cause termination of transcription and addition of polyadenylated
XX      CC      ribonucleotides to a 3' end of the mRNA molecule). The ESTs are useful as
XX      CC      molecular tool for the targeting and isolation of novel genes for plant
XX      CC      protection and improvement. The ESTs are useful for developing new
XX      CC      strategies for understanding critical plant developmental and metabolic
XX      CC      pathways, for isolating genes and promoters, for identifying and mapping
XX      CC      the genes involved in developmental and metabolic pathways, and for
XX      CC      determining gene function. The cotton nucleic acid molecules are useful
XX      CC      as molecular tags to isolate genetic regions, isolate genes, map genes,
XX      CC      and determine gene function. The nucleic acid molecules are useful for
XX      CC      determining if genes are members of a particular gene family and for use
XX      CC      in marker-assisted breeding programs. The present sequence is one of the
XX      CC      4930 cotton ESTs of the invention. NOTE: The sequences are not displayed
XX      CC      in the specification but are available in electronic format from the
XX      CC      USPTO at seqdata.uspto.gov/sequence.html?DocID=20040123338.
XX
XX      SQ      Sequence 377 BP; 100 A; 87 C; 90 G; 100 T; 0 U; 0 Other;
XX
Query Match      1.6%; Score 29; DB 12; Length 377;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      1233 GACAGCAAGAGAGCCATGATGATGATCAC 1261
XX      |||||||
XX      DB      161 GACAGCAAGAGAGCCATGATGATGATCAC 189

```

Search completed: June 8, 2005, 07:26:59
Job time : 970.851 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:49:46 ; Search time 5995.12 Seconds

(without alignments)
11174.613 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctcttcttctcttcttcttgc.....aattgtagtataatttcttgc 1760

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	750	42.6	813	7	CK768601 Gm-r1030-
2	552	31.9	686	2	CK768601 Gm-r1030-
3	533	30.3	533	4	BM523576 Bam85904.
4	527	29.9	672	2	AM508269 s151n10.Y
5	517	29.4	622	2	AM397496 sg79e12.Y
6	508	28.9	619	2	BE313163 sg98f01.Y
7	505	28.7	575	2	AM397453 sg79d01.Y
8	501	28.5	648	1	AT794750 sg58f07.Y
9	475	27.0	539	7	AM317927 sg58f08.Y
10	469	26.6	459	7	CK769092 Gm-r1030-
11	467	26.5	478	7	CK769092 Gm-r1030-
12	443	25.2	443	5	BO612082 sg27b07.Y
13	400	22.7	420	4	BI347339 P165E19.C
14	391	22.2	485	2	BE609839 sg45a10.Y
15	383	21.8	435	2	AM472457 sg12d05.Y
16	375	21.3	496	2	AM509380 sg12d05.Y
17	363	20.6	451	4	BG726672 sg23b08.Y
18	359	20.4	576	1	AI496426 sg04e09.Y
19	315	17.9	316	2	AM398011 sg71b06.Y
20	289	17.0	308	2	AM705757 sg51e05.Y
21	280	16.5	290	2	AM460108 sg11d02.Y
22	279	15.9	408	5	BU763677 sg84e09.Y
23	276	15.7	661	2	BE191464 sg76h10.Y
24	246	14.0	552	4	BM308082 sg41f07.Y

25	243	13.8	396	2	AM472088	44	AM472088 s119g09.Y
26	205	11.6	208	2	AM569076	45	AM569076 s163b05.Y
27	201	11.4	201	4	BG404951		BG404951 sg46d03.Y
28	172	9.8	271	2	AM568795		AM568795 s161b10.Y
29	152	8.6	152	4	BG237745		BG237745 sg17a01.Y
30	152	8.6	152	4	BG651312		BG651312 sg8d3a07.Y
31	146	8.3	304	2	BP425513		BP425513 sg55f04.Y
32	146	8.3	449	2	AM203517		AM203517 sg135b05.Y
33	146	8.3	516	4	BG652636		BG652636 sg46g02.Y
34	146	8.3	540	5	BM887128		BM887128 sg34e09.Y
35	146	8.3	552	2	BM955039		BM955039 sg34e09.Y
36	146	8.3	634	2	BE313050		BE313050 sg03e10.Y
37	145	8.2	670	4	BG044525		BG044525 sg22e07.Y
38	143	8.1	563	2	AM472315		AM472315 sg123f04.Y
39	138	7.8	158	2	BE603020		BE603020 sg29gmaxSC
40	135	7.7	388	2	AM432728		AM432728 sg85f08.Y
41	135	7.7	411	2	AM099866		AM099866 sg17e12.Y
42	135	7.7	420	4	BM092617		BM092617 sg16d10.Y
43	135	7.7	456	2	AM620996		AM620996 sg150e08.Y
44	135	7.7	457	2	AM310359		AM310359 sg135b05.X
45	135	7.7	467	1	AI442850		AI442850 sg27f05.X

ALIGNMENTS

RESULT 1
CK768601
LOCUS Gm-r1030-621 Gm-r1030 Glycine max cDNA clone Gm-r1030-621 5', mRNA
DEFINITION sequence.
ACCESSION CK768601
VERSION CK768601.1 GI:42722702
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE

1 (bases 1 to 813)
Periapuram, C.C., Li, L., Wurtel, E.S., Westgate, M.E. and Nikolau, B.J.
Molecular determination of soybean composition
Unpublished (2004)
Contact: Basil J. Nikolau and Cyril C. Periapuram
Department of Biochemistry, Biophysics and Molecular Biology
Iowa State University
2210 Molecular Biology Building, Ames, IA 50011, USA
Tel: 515 294 9423
Fax: 515 294 0453
Email: dimmas@iastate.edu

TITLE

Individual base call and confidence value were assigned using the
Phred software (http://www.phrap.org/). Overall sequence quality
assessment and vector trimming were conducted using the Lucy
software (http://www.tigr.org/software/).
This clone was originally generated by the Public Soybean EST
Project (http://129.186.26.94/soybeanest.html)/Shoemaker, R
(rcshoe@iastate.edu).
This clone is available through: Biogenetic Services, 801 32nd Ave,
Brookings, SD 57006. For further information call 605-697-8500 or
contact info@biogeneticservices.com
Seq primer: T7-1 (5' AAT AGC ACT CAC TAT AG 3').
Location/Qualifiers

FEATURES

source
1..813
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="Gm-r1030-621"
/lab_host="DH10B"
/clone_1lb="Gm-r1030"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from

immature cotyledons of greenhouse grown plants (individual seed fresh weight of 100-300mg). The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note that Gm-r1021 is a re-rack of Gm-cl007."

ORIGIN

Query Match 42.6%; Score 750; DB 7; Length 813;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 800; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

487 GCACATGAACTGGCTGATGCGCCAGGCAAGGCTTTGACATGATTGACAGA 546
 13 GCACATGAACTGGCTGATGCGCCAGGCAAGGCTTTGACATGATTGACAGA 72

547 AGCAGTTGAGGCTTACATGAAATCCATCTTCCCGAATCTATGACCCGATT 606
 73 AGCAGTTGAGGCTTACATGAAATCCATCTTCCCGAATCTATGACCCGATT 132

607 TCATTGCTCCCAACCAAGAGAGCGGCCAACAGCTCATCAAGGCCAACAAAGCAG 666
 133 TCATTGCTCCCAACCAAGAGAGCGGCCAACAGCTCATCAAGGCCAACAAAGCAG 192

667 AAGTTCAACAATTCATCAAGACATCAAGAGCGCTTTAAGAGACCAACAAAGTGA 726
 193 AAGTTCAACAATTCATCAAGACATCAAGAGCGCTTTAAGAGACCAACAAAGTGA 252

727 TGGTGTACTGTGAGACTGCCAACACAGAGAGGTACAGTAATTTGTTGGGCTTATG 786
 253 TGGTGTACTGTGAGACTGCCAACACAGAGAGGTACAGTAATTTGTTGGGCTTATG 312

787 AACACATGAGAAATCTTTGGCTGCTGTGTGACAGAAATAGAGCTGATTTCTCCCTCA 846
 313 AACACATGAGAAATCTTTGGCTGCTGTGTGACAGAAATAGAGCTGATTTCTCCCTCA 372

847 CCTTGATGCCATGCTTGTGTATGAAATGTTCTTCAATTAATGAGACCCCTCAGA 906
 373 CCTTGATGCCATGCTTGTGTATGAAATGTTCTTCAATTAATGAGACCCCTCAGA 432

907 AACCTTTGTACAGGAGGCTGATTTGATCTTGCCATGCGAGAACTTTGATTTGAG 966
 433 AACCTTTGTACAGGAGGCTGATTTGATCTTGCCATGCGAGAACTTTGATTTGAG 492

967 ATGACTTCAAGAGTGTGACCAAAATGAAATGTTGTTGTTGTTCTTTGTTGGGG 1026
 493 ATGACTTCAAGAGTGTGACCAAAATGAAATGTTGTTGTTGTTCTTTGTTGGGG 552

1027 CTGATCAAGCAACATCTATAGTCAATCAACATCTGGGAAACAAATGATGATGA 1086
 553 CTGATCAAGCAACATCTATAGTCAATCAACATCTGGGAAACAAATGATGATGA 612

1087 ATCTTTGGCTCCCAAACTTTCCGTTCCAAAGAAATCTCCAAAGCAACGTTTGTATG 1146
 613 ATCTTTGGCTCCCAAACTTTCCGTTCCAAAGAAATCTCCAAAGCAACGTTTGTATG 672

1147 AATAGTCAACAGCAATGTCATCTTATAGAGCTGTGTAACATCCAGCATGTTGTTG 1206
 673 AATAGTCAACAGCAATGTCATCTTATAGAGCTGTGTAACATCCAGCATGTTGTTG 732

1207 TTATTAGATGAGCTTACGTAGGAGACAGCAAGAGCAATGATGATGACATTTGAG 1266
 733 TTATTAGATGAGCTTACGTAGGAGACAGCAAGAGCAATGATGATGACATTTGAG 792

1267 AGATATTCATGGGTGAAAGA 1287

Db 793 AGATATTCATGGGTGAAAGA 813

RESULT 2
 LOCUS AM348857/c 686 bp mRNA linear EST 04-OCT-2000
 DEFINITION GM210010A10E12 Gm-r1021 Glycine max cDNA clone Gm-r1021-3671 3',
 mRNA sequence.
 ACCESSION AM348857
 VERSION AM348857.1 GI:6846567
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosoids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseolae;
 Glycine.

1 (bases 1 to 686)
 Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,
 Erpelting, J., Raph, C., Shoop, E., Pardinas, V., Liu, L. and Lewin, H.
 A Functional Genomics Program for Soybean (NSF 9872565)
 Unpublished (1999)
 Other ESTs: A1496426
 Contact: Vodkin, L.O., PI, A Functional Genomics Program for
 Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional
 Genomics

University of Illinois
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l-vodkin@uiuc.edu
 This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134. For further information
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or info@genome
 systems.com web site: www.genomesystems.com
 Seq primer: 5'-TCTTTTCTTTTCTTTTCTTTT(A/C/G)-3'.

FEATURES

source
 1..686
 /organism="Glycine max"
 /mol_type="mRNA"
 /cui="var="Williams"
 /db_xref="taxon:3847"
 /clone="Gm-r1021-3671"
 /tissue_type="root"
 /lab_host="XLL10-Gold"
 /clone_lib="Gm-r1021"
 /note="Vector: pBluescript II XR; Site 1: EcoRI, Site 2:
 XhoI; Library Gm-r1021 is a sequence-driven, re-racked set
 of the original library Gm-cl004 which was prepared from
 root cDNA. The mRNA was isolated from entire roots of 8
 day old 'Williams' seedlings which were propagated on
 paper towels with distilled water. StrataGene's cDNA
 Synthesis Kit (catalog #200401) was used to synthesize
 the cDNA. The Gm-cl004 library was constructed by Dr.
 Paul Keim & Virginia H. Coryell, Department of Biology,
 Box5640, Northern Arizona University, Flagstaff, AZ
 86011, email: paul.keim@uau.edu, virginia.coryell@uau.edu.
 The contig analysis to select unique genes was performed
 by the laboratory of Ernest Retzel, Computational Biology
 Centers, University of Minnesota,
 http://www.cbc.umn.edu/Research/Projects/Soybean/index.html
 Retacking was performed by Genome Systems, St. Louis,
 http://www.genomesystems.com, and sequencing by the Keck
 Center for Comparative and Functional Genomics,
 University of Illinois,
 http://www.life.uiuc.edu/biotech/keck.html."

ORIGIN

Query Match 31.9%; Score 562; DB 2; Length 686;
 Best Local Similarity 100.0%; Pred. No. 5.2e-305; Indels 0; Gaps 0;
 Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
AM508269	672 bp	mRNA linear EST 03-DEC-2001	AM508269									
<p>g515h10.y1 Gm-r1030 Glycine max cDNA clone GENOME EST ID: Gm-r1030-2492.5' similar to SW:INOL.CITPA.P42802</p> <p>MYO-INOSITOL-1-PHOSPHATE SYNTHASE ; mRNA sequence.</p> <p>AM508269</p> <p>AM508269.1 GI:7146347</p> <p>EST.</p> <p>Glycine max (soybean)</p> <p>Glycine max</p> <p>Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.</p> <p>1 (bases 1 to 672)</p> <p>Shoemaker,R., Keim,P., Vodkin,L., Expelding,J., Coryell,V., Khanna,A., Bolla,B., Marras,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Pearson,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Rutter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.</p> <p>Public Soybean EST Project</p> <p>Unpublished (1999)</p> <p>Contact: Shoemaker R/Public Soybean EST Project</p>												

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp., 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: cs@resgen.com
Insert Length: 1864 Std Error: 0.00
High quality sequence stop: 431.
Location/Qualifiers
1..672

```

1. .672
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-r1030-2492"
/lab_host="DH10B"
/clone_lib="Gm-r1030"
/notice="vector: pSPORT1, Site_1: SalI, Site_2: NotI, This
cDNA library was constructed from mRNA isolated from
immature cotyledons of greenhouse grown plants
(individual seed fresh weight of 100-300mg). The library
was prepared using the Life Technologies pSuperScript cDNA
library construction kit. Complementary DNA was
synthesized from mRNA using a poly(dT) sequence with a
NotI restriction site. SalI linkers adapters were ligated
to the blunt-ended cDNA fragments followed by NotI
digestion. The cDNA fragments were directionally cloned
into the NotI-SalI restriction site of the pSPORT1
vector. The ligated cDNA fragments were transformed into
E. coli Electromax DH10B host cells. This library was
constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note
that Gm-r1030 is a re-track of Gm-cl007."

```

ORIGIN

	Query Match	29.9%	Score 527	DB 2;	Length 672	
	Best Local Similarity	100.0%	Pred. No. 3.1e-285			
	Matches	527	Conservative	0	Mismatches	0; Indels 0; Gaps 0.
Oy	18	TGTAATTCATTCACTTCTAATCCTTGGAAGAAATAATGTTCATGAGAAATTTAAGGT	77			
	[]	[]	[]	[]	[]	[]
Dd	1	TGTAATTCATTCACTTCTAATCCTTGGAAGAAATAATGTTCATGAGAAATTTAAGGT	60			
Oy	78	GAGGTCCCTAATGTGAGACACCGAGACTGAGTTCAGTCCGTGACAACACTAGAAAACC	137			
	[]	[]	[]	[]	[]	[]
Dd	61	GAGTCTCCTAATGTGAGACACCGAGACTGAGTTCAGTCCGTGACAACACTAGAAAACC	120			
	[]	[]	[]	[]	[]	[]

Qy	138	ACCGAATCTGTTCAACGGAACAGGAATGGCACTATCAGTGAATGTCAAAACCCAAATCT	197
Db	121	ACCGAATCTGTTCAACGGAACAGGAATGGCACTATCAGTGAATGTCAAAACCCAAATCT	180
Qy	198	GTCAAAATACGAATTTTAAAAACCAACATCCATGTTCTAAATTAGGGATATGCTTGGGT	257
Db	181	GTCAAAATACGAATTTTAAAAACCAACATCCATGTTCTAAATTAGGGATATGCTTGGGT	240
Qy	258	TGGGGTGAAAACAACGGCTCAACCTCACCGGTGTATTATTCGTAACCGAGAGGGCAATT	317
Db	241	TGGGGTGAAAACAACGGCTCAACCTCACCGGTGTATTATTCGTAACCGAGAGGGCAATT	300
Qy	318	TCATGGGGCTACAAAGACAAATTCACAAACGAATTACTTTGGCTCCCTCACCCAAGCC	377
Db	301	TCATGGGGCTACAAAGACAAATTCACAAACGAATTACTTTGGCTCCCTCACCCAAGCC	360
Qy	378	TCAGCTATCCGAGTGGGTCTCTCCACAGGAGAGAAATCTATGCCCCATTCAAGAGCCTG	437
Db	361	TCAGCTATCCGAGTGGGTCTCTCCACAGGAGAGAAATCTATGCCCCATTCAAGAGCCTG	420
Qy	438	CTTCCAAATGTTAAACCTGACGACATTGTGTTTGGGGGATGGGAATATCAGCAACATGAAC	497
Db	421	CTTCCAAATGTTAAACCTGACGACATTGTGTTTGGGGGATGGGAATATCAGCAACATGAAC	480
Qy	498	CTGGCTGATGCATGAGCCACAGGCAAAAGGTGTTTGACATCGAATTTGCA 544	
Db	481	CTGGCTGATGCATGAGCCACAGGCAAAAGGTGTTTGACATCGAATTTGCA 527	

RESULT	5
LOCUS	AM397496
DEFINITION	AM397496 622 bp mRNA linear EST 14-JUL-2004
ACCESSION	AF099426 Gm-G1007 Glycine max cDNA clone GENOME SYSTEMS CLONE ID
VERSION	Gm-G1007-2663 5' similar to SM:INOL CITPA P42802
KEYWORDS	MNO-INOSITOL-1-PHOSPHATE SYNTHASE ;, mRNA sequence.
SOURCE	AM397496.1 GI:691596
ORGANISM	EST. Glycine max (soybean) Glycine max

REFERENCE	TITLE
<p> <i>Authors</i> 1 (bases 1 to 622) Shoemaker, R., Kelm, P., Vodka, L., Expelling, J., Corryell, V., Khanna, A., Bolla, B., Marras, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, R., Steptoe, M., Treising, B., Allen, M., Bowers, Y., Pearson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R. Public Soybean EST Project Unpublished (1999) Contact: Shoemaker R/PUBLIC Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel.: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone) 605 423 4165; email: info@biogeneticservices.com Seq primer: -40RP from Gibco High quality sequence stop: 434. </p>	<p> <i>Journal</i> Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine. </p>

FEATURES	location/Qualifiers
source	1. .622
	/organism="Glycine max"
	/mol_type="mRNA"
	/cultivar="Williams"
	/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1007-2663"
/lab_host="DH10B"
/clone_lib="Gm-c1007"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from
immature cotyledons of greenhouse grown plants (individual
seed fresh weight of 100-300mg). The library was prepared
using the Life Technologies pSuperscript cDNA library
construction kit. Complementary DNA was synthesized from
mRNA using a poly (dT) sequence with a NotI restriction
site. SalI linkers adapters were ligated to the
blunt-ended cDNA fragments followed by NotI digestion. The
cDNA fragments were directionally cloned into the
NotI-SalI restriction site of the pSPORT1 vector. The
ligated cDNA fragments were transformed into E. coli
Electromax DH10B host cells. This library was constructed
by Dr. Lila Vodka and Dr. Anu Khanna."

ORIGIN

Query Match 29.4%; Score 517; DB 2; Length 622;
Best Local Similarity 99.8%; Pred. No. 1.3e-279;
Matches 567; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 487 GCACATGAACCTGGCTGATGCCAGGCAAGGTTTGACATGATTTGACAG 546
DB 18 GCACATGAACCTGGCTGATGCCAGGCAAGGTTTGACATGATTTGACAG 77
QY 547 AGCAGTTGAGGCTTATCATGGAATCCATGCTTCCCGGATCTATGACCCGGATT 606
DB 78 AGCAGTTGAGGCTTATCATGGAATCCATGCTTCCCGGATCTATGACCCGGATT 137
QY 607 TCACTTCTGCCAACCAAGAGAGGCTGCCAACCACTCATCAAGGCAACCAAGAGC 666
DB 138 TCACTTCTGCCAACCAAGAGAGGCTGCCAACCACTCATCAAGGCAACCAAGAGC 197
QY 667 AAGTTCAACCAATCATCAAGAGAGGCTGCCAACCACTCATCAAGGCAACCAAGAG 726
DB 198 AAGTTCAACCAATCATCAAGAGAGGCTGCCAACCACTCATCAAGGCAACCAAGAG 257
QY 727 TGGTTGACTGTGAGTCCCAACAGAGAGTACAGTAATTTGTTGGGCTTAAATG 786
DB 258 TGGTTGACTGTGAGTCCCAACAGAGAGTACAGTAATTTGTTGGGCTTAAATG 317
QY 787 AACCAATGAGAACTCTTGGCTGCTGTGACAGAAATGAGGCTGATTTCTCTTCCA 846
DB 318 AACCAATGAGAACTCTTGGCTGCTGTGACAGAAATGAGGCTGATTTCTCTTCCA 377
QY 847 CCTTGTATGCAATGCTGTGTTGTAAGAAATGTCCTTTCATTATGGAAGCCTCAG 906
DB 378 CCTTGTATGCAATGCTGTGTTGTAAGAAATGTCCTTTCATTATGGAAGCCTCAG 437
QY 907 ACACTTTTGTACAGGAGCTGATTTGATCTTGCCATCGCAGAGAACCTTGTATGCTGAG 966
DB 438 ACACTTTTGTACAGGAGCTGATTTGATCTTGCCATCGCAGAGAACCTTGTATGCTGAG 497
QY 967 ATGACTTTCAAGATGCTGACAGCAAAATGAAATCTGTGTTGATTTCTTGTGGGG 1026
DB 498 ATGACTTTCAAGATGCTGACAGCAAAATGAAATCTGTGTTGATTTCTTGTGGGG 557
QY 1027 CTGTATCAAGCAACATCTATAGTCAG 1054
DB 558 CTGTATCAAGCAACATCTATAGTCAG 585

RESULT 6
BE331363/c
LOCUS
DEFINITION
Gm-c1041-1394 bp mRNA linear EST 13-JUL-2004
8098101.y1 Gm-c1041 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
1-PHOSPHATE SYNTHASE ; mRNA sequence.
ACCESSION
VERSION
KEYWORDS
BE331363.1 GI:9205139
EST.

SOURCE
ORGANISM
Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
1 (bases 1 to 619)

AUTHORS

Shoemaker, R., Keim, P., Vodka, L., Erpelting, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N.,
McGarr, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McGarr, R., Waterston, R. and Wilson, R.

TITLE

Public Soybean EST Project

JOURNAL

Unpublished (1999)

COMMENT

Contact: Shoemaker R./Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. Possible
reversed clone: similarity on wrong strand This clone is available
through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA
(phone: 800 423 4163; email: info@biogeneticservices.com)
High quality sequence stop: 491.

FEATURES

source

1..619

/organism="Glycine max"

/mol_type="mRNA"

/cultivar="Williams"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1041-1394"

/tissue_type="Senescing leaves, mature plant, greenhouse
grown"

/lab_host="DH10B"

/clone_lib="Gm-c1041"

/note="Vector: pT73Pac (Pharmacia); Site 1: EcoRI;
Site 2: HindIII; This library was constructed from mRNA
isolated from senescing leaf tissue of mature greenhouse
grown plants of the cultivar Williams. Complementary DNA
was synthesized from mRNA using a 3' anchored poly(dT)
primer. EcoRI adapters were ligated to the blunt-ended
cDNA fragments followed by digestion with EcoRI and
HindIII. The cDNA fragments were directionally cloned
into the EcoRI-HindIII restriction site of the pT73-Pac
vector. The ligated cDNA fragments were transformed into
DH10B host cells. This library was constructed by Dr.
Randy Shoemaker."

ORIGIN

Query Match 28.9%; Score 508; DB 2; Length 619;
Best Local Similarity 100.0%; Pred. No. 1.6e-274;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1252 ATGAGTACACTTCAGAGATATTCATGGGTGAAGAGCACTGTTTGCACAAACAT 1311
DB 519 ATGAGTACACTTCAGAGATATTCATGGGTGAAGAGCACTGTTTGCACAAACAT 460
QY 1312 GCGAGGATTCCTCTTAGCTGCTCTTATATCTTGACACTGGTCTTCTGCTGAGCTCA 1371
DB 459 GCGAGGATTCCTCTTAGCTGCTCTTATATCTTGACACTGGTCTTCTGCTGAGCTCA 400
QY 1372 GCACTAGAAATGAGTTAAAGCTGAAATGAGGAAATTCACATCATTCACCCAGTTG 1431
DB 399 GCACTAGAAATGAGTTAAAGCTGAAATGAGGAAATTCACATCATTCACCCAGTTG 340
QY 1432 CTACCATCTCAGCTACCTCAACAGAGCTCTCTGTTTCAACCGGGTACACCAAGTGTGA 1491
DB 339 CTACCATCTCAGCTACCTCAACAGAGCTCTCTGTTTCAACCGGGTACACCAAGTGTGA 280

```
QY 1492 ATGCAATGTCACAGCAGCGTCGCAATGCTGGAAGAAACATAATGAGCGCTTGTTGGATTGG 1551
DB 279 ATGCAATGTCACAGCAGCGTCGCAATGCTGGAAGAAACATAATGAGCGCTTGTTGGATTGG 220
QY 1552 CCCCAGAGAAATTAACATGATTTCTCGAGTACAAAGTGAAGCATGGAGCCGGAAGATAATATG 1611
DB 219 CCCCAGAGAAATTAACATGATTTCTCGAGTACAAAGTGAAGCATGGAGCCGGAAGATAATATG 160
QY 1612 TTGGGGTACCTAGCTGATGATTTTATGTTTATTAATATATGCTTATTAATTTTGGCAAGT 1671
DB 159 TTGGGGTACCTAGCTGATGATTTTATGTTTATTAATATATGCTTATTAATTTTGGCAAGT 100
QY 1672 GTAATTAATGATGATGATGCTTCAATTAATGCTTTAGAGCGGGCAATATCTGTTACTAGGA 1731
DB 99 GTAATTAATGATGATGATGCTTCAATTAATGCTTTAGAGCGGGCAATATCTGTTACTAGGA 40
QY 1732 ACATGAATGAATGATGATTAATTTGTG 1759
DB 39 ACATGAATGAATGATGATTAATTTGTG 12

RESULT 7
AM397453 575 bp mRNA linear EST 14-JUL-2004
LOCUS 8973b01.y1 Gm-c1007 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-c1007-2594 5' similar to SW:IN01_CITPA P42802
ACCESSION MYO-INOSITOL-1-PHOSPHATE SYNTHASE ;, mRNA sequence.
VERSION AM397453
KEYWORDS AM397453.1 GI:6915923
SOURCE EST.
ORGANISM Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 575)
REFERENCE Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
AUTHORS Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCaun,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the "Other ESTs on clone" field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Insert Length: 1909 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 414.
Location/Qualifiers
1. 575
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1007-2594"
/lab_host="DH10B"
/clone_lib="Gm-c1007"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from
immature cotyledons of greenhouse grown plants (individual
seed fresh weight of 100-300mg). The library was prepared
using the Life Technologies pSuperScript cDNA library
```

construction kit. Complementary DNA was synthesized from mRNA using a poly (dt) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

```
ORIGIN
Query Match 28.7%; Score 505; DB 2; Length 575;
Best Local Similarity 100.0%; Pred. No. 8e-273;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 CTTTGTAAATTTCTATTCATTTCTTATGTTGTAAGAAATATGTTCAAGAAATTTTA 72
DB 1 CTTTGTAAATTTCTATTCATTTCTTATGTTGTAAGAAATATGTTCAAGAAATTTTA 60
QY 73 AGTTGAATGTCCTAATGATGATGACCGAGACTGATTCAGTCCGTGACACTAGC 132
DB 61 AGTTGAATGTCCTAATGATGATGACCGAGACTGATTCAGTCCGTGACACTAGC 120
QY 133 AAACCAACCAACTGTTTACAGAGAAACAGAAATGCACTATCAGTGTGCAAAACCA 192
DB 121 AAACCAACCAACTGTTTACAGAGAAACAGAAATGCACTATCAGTGTGCAAAACCA 180
QY 193 AATCTGTCAAAATACGAATTTTAAACCAACATCCATGTTCTTAATTAAGGGTAAATGCTTG 252
DB 181 AATCTGTCAAAATACGAATTTTAAACCAACATCCATGTTCTTAATTAAGGGTAAATGCTTG 240
QY 253 TGGGTTGGGGTGGAAACACAGCGCTCAACCTCAGCGGTGTGTTATGTTCTTACCGAGAG 312
DB 241 TGGGTTGGGGTGGAAACACAGCGCTCAACCTCAGCGGTGTGTTATGTTCTTACCGAGAG 300
QY 313 GCATTTCAATGGGCTACAAAGAACAGATTCAACAGCCAAATTAATCTTGGTCCCTCACCC 372
DB 301 GCATTTCAATGGGCTACAAAGAACAGATTCAACAGCCAAATTAATCTTGGTCCCTCACCC 360
QY 373 AAGCCTCAGCTATCCGAGTGGGTCTCTCCAGGGAGAGAAATCATGCCCCCATTCAGA 432
DB 361 AAGCCTCAGCTATCCGAGTGGGTCTCTCCAGGGAGAGAAATCATGCCCCCATTCAGA 420
QY 433 GCGTCTCTCCAAATGGTTAAACCTGACGACATTTGTTGGGGAGATGAGATATCAGACA 492
DB 421 GCGTCTCTCCAAATGGTTAAACCTGACGACATTTGTTGGGGAGATGAGATATCAGACA 480
QY 493 TGAACCTGGCTGATGCGATGGCCAG 517
DB 481 TGAACCTGGCTGATGCGATGGCCAG 505

RESULT 8
AT194750 648 bp mRNA linear EST 12-JUL-2004
LOCUS sb66f07.y1 Gm-c1019 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-c1019-158 5' similar to SW:IN01_CITPA P42802
ACCESSION MYO-INOSITOL-1-PHOSPHATE SYNTHASE ;, mRNA sequence.
VERSION AT194750
KEYWORDS AT194750.1 GI:5342466
SOURCE EST.
ORGANISM Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 648)
REFERENCE Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
AUTHORS Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
```

TITLE
JOURNAL
COMMENT

McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewaterston.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. Possible
reversed clones: similarity on wrong strand this clone is available
through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA
(phone: 800 423 4163; email: info@biogeneticservices.com)
Insert Length: 1867 Std Error: 0.00
High quality sequence stop: 438.
Location/Qualifiers

FEATURES
source

1..648
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl019-158"
/issue_type="Immature seed coats of greenhouse grown
plants"
/lab_host="DH10B (Gibco BRL)"
/clone_lib="Gm-cl019"
/note="Vector: pSPORT1 (Life Technologies); Site 1: Not I;
Site 2: Sal I; This cDNA library was constructed from mRNA
isolated from immature seed coats (200-300 mg) of
greenhouse grown plants. The library was prepared using
the Life Technologies pSuperScript cDNA library
construction kit. Complementary DNA was synthesized from
mRNA using a poly (dT) sequence with a Not I restriction
site. Sal I linker adapters were ligated to the
blunt-ended cDNA fragments followed by Not I digestion.
The cDNA fragments were directionally cloned into the Not
I-Sal I restriction site of the pSPORT1 vector. The
ligated cDNA fragments were transformed into E.coli
Electromax DH10B host cells (Gibco BRL). This library was
constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

ORIGIN

Query Match 28.5%; Score 501; DB 1; Length 648;
Best Local Similarity 99.8%; Pred. No. 1.5e-270;
Matches 551; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 38 ATCTTGTAAATAATGTTTCATCGAATTTTAACTGCTGCTTAATGTGAAGTA 97
DB 18 ATCTTGTAAATAATGTTTCATCGAATTTTAACTGCTGCTTAATGTGAAGTA 77
QY 98 CACCGAGCTGAGATTCACTCCGCTGACAACTACGAAACACCGAATCTTTCAGCAGAA 157
DB 78 CACCGAGCTGAGATTCACTCCGCTGACAACTACGAAACACCGAATCTTTCAGCAGAA 137
QY 158 CAGGAATGACCTTCACTGATGATGTCAAAACCAATCTGTCAATATGCAATTTTAAAC 217
DB 138 CAGGAATGACCTTCACTGATGATGTCAAAACCAATCTGTCAATATGCAATTTTAAAC 197
QY 218 CAAATCCTATGTTTCTTAATTAAGGGTATGCTTGGGTTGGGTTGGAACAACAGGCTC 277
DB 198 CAAATCCTATGTTTCTTAATTAAGGGTATGCTTGGGTTGGGTTGGAACAACAGGCTC 257
QY 278 AACCTTCAACCGATGATGTTTATGCTAAGCAGAGGAGCATTTTCATGGGCTACAAAGACAA 337
DB 258 AACCTTCAACCGATGATGTTTATGCTAAGCAGAGGAGCATTTTCATGGGCTACAAAGACAA 317
QY 338 GATTCAACAAGCAATTAATCTTGGCTCCTACCCAGGCTCAGCTATCCGAGTTGGGTC 397
DB 318 GATTCAACAAGCAATTAATCTTGGCTCCTACCCAGGCTCAGCTATCCGAGTTGGGTC 377
QY 398 CTTCCAGGAGGAGGAATCTATGCCCATTCGAAGACCTGCTTCCAAATGTTAAACCTGA 457

DB 378 CTTCCAGGAGGAGGAATCTATGCCCATTCGAAGACCTGCTTCCAAATGTTAAACCTGA 437
QY 458 CCAATCTGTTTGGGGGATGAGATATCAGCAACATGAACCTGCTGATGATCCATGGCCAG 517
DB 438 CCAATCTGTTTGGGGGATGAGATATCAGCAACATGAACCTGCTGATGATCCATGGCCAG 497
QY 518 GCGAAGCTGTTTGAATCATGATTTGAGAGAGAGATTTAGAGCCCTTACATGAATTCATGCT 577
DB 498 NCCAAAGGTTTGAATCATGATTTGAGAGAGAGATTTAGAGCCCTTACATGAATTCATGCT 557
QY 578 TCACCTCCCGG 589
DB 558 TCACCTCCCGG 569

RESULT 9
AM317927
LOCUS
DEFINITION
G558h08.y1 Gm-cl007 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl007-664 5' similar to FR:040271 Q40271
MYO-INOSITOL-1-PHOSPHATE SYNTHASE. ; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AM317927
AM317927.1 GI:6747471
EST.
Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
AUTHORS
1 (baaes 1 to 539)
Shoemaker, R., Keim, P., Vodkin, L., Expelding, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Persson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N.,
Schurk, R., Riter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewaterston.wustl.edu

When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 432.
Location/Qualifiers

FEATURES
source

1..539
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl007-664"
/lab_host="DH10B"
/clone_lib="Gm-cl007"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from
immature cotyledons of greenhouse grown plants (individual
seed fresh weight of 100-300mg). The library was prepared
using the Life Technologies pSuperScript cDNA library
construction kit. Complementary DNA was synthesized from
mRNA using a poly (dT) sequence with a NotI restriction
site. SalI linker adapters were ligated to the
blunt-ended cDNA fragments followed by NotI digestion. The
cDNA fragments were directionally cloned into the
NotI-SalI restriction site of the pSPORT1 vector. The

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

REFERENCE 1 (bases 1 to 478)

AUTHORS Perlapuram, C.C., Li, L., Murtelle, E.S., Westgate, M.E. and Nikolau, B.J.

TITLE Molecular determination of soybean composition

JOURNAL Unpublished (2004)

COMMENT Contact: Basil J. Nikolau and Cyril C. Perlapuram

Department of Biochemistry, Biophysics and Molecular Biology
Iowa State University
2210 Molecular Biology Building, Ames, IA 50011, USA

Tel: 515 294 9423

Fax: 515 294 0453

Email: dimmas@iastate.edu

Individual base call and confidence value were assigned using the Phred software (<http://www.phrap.org/>). Overall sequence quality assessment and vector trimming were conducted using the Lucy software (<http://www.tigr.org/software/>).

This clone was originally generated by the Public Soybean EST Project (<http://129.186.26.94/soybeanest.html>) / Shoemaker, R (rshoemaker@iastate.edu).

This clone is available through: Biogenetic Services, 801 32nd Ave, Brookings, SD 57006. For further information call 605-697-8500 or contact info@biogeneticservices.com

Seq primer: T7-1 (5' AAT ACG ACT CAC TAT AG 3').

Location/Qualifiers

1..478

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="Gm-r1030-2818"

/lab_host="DH10B"

/clone_1lb="Gm-r1030"

/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This cDNA library was constructed from mRNA isolated from immature cotyledons of greenhouse grown plants

(individual seed fresh weight of 100-300mg). The library was prepared using the Life Technologies Superscript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a

NotI restriction site. SalI linker adapters were ligated to the blunt-ended cDNA fragments followed by NotI

digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1

vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. This library was

constructed by Dr. Lila Vodkin and Dr. Ann Khanna. Note

that Gm-r1030 is a re-rack of Gm-cl007."

ORIGIN

Query Match 26.5%; Score 467; DB 7; Length 478;
Best Local Similarity 100.0%; Pred. No. 2.3e-251;

Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 19 GGAATTCATTCATTTTATCTTTGTGAAAAATATGTTTCATCGAATTTTAAAGTTG 78
Db 12 GGAATTCATTCATTTTATCTTTGTGAAAAATATGTTTCATCGAATTTTAAAGTTG 71
Oy 79 AGTGTCTTATGTAAGTACACGAGATGAGTCCGCTGTAACAATCGAAGACA 138
Db 72 AGTGTCTTATGTAAGTACACGAGATGAGTCCGCTGTAACAATCGAAGACA 131
Oy 139 CGGAATCTGTCAGGAGACAGGAATGCACTTATCAGTGAATGTCMAAACCAGATCTG 198
Db 133 CGGAATCTGTCAGGAGACAGGAATGCACTTATCAGTGAATGTCMAAACCAGATCTG 191
Oy 199 TCAATATCAATTTAAACCAACATCATGTTCTTAATTTAGGGGTAATGCTTGGGTT 258
Db 192 TCAATATCAATTTAAACCAACATCATGTTCTTAATTTAGGGGTAATGCTTGGGTT 251
Oy 259 GGGGTGAAACAACGAGCTCAACCGGTGGTGTATGCTAAACGAGAGGAGCATTT 318

```

```

Db 252 GGGGTGAAACAACGAGCTCAACCGGTGGTGTATGCTAAACGAGAGGAGCATTT 311
Oy 319 CATGGCTACAAAGGACAAAGATTGCAACAGCCATTTAGCTCCCTCAACCAAGCTT 378
Db 312 CATGGCTACAAAGGACAAAGATTGCAACAGCCATTTAGCTCCCTCAACCAAGCTT 371
Oy 379 CAGCTATCCGAGTGTGGTCTCTTCAGAGAGAGAAATCATGCCCCATTCAGAGCTGC 438
Db 372 CAGCTATCCGAGTGTGGTCTCTTCAGAGAGAGAAATCATGCCCCATTCAGAGCTGC 431
Oy 439 TTCCAAATGTTAAACCTGACGACATTTGTTGGGGAGATGGATATC 485
Db 432 TTCCAAATGTTAAACCTGACGACATTTGTTGGGGAGATGGATATC 478

```

RESULT 12
BO612082
LOCUS
DEFINITION
BO612082 443 bp mRNA linear EST 05-JUL-2004
sap7b07.y1 Gm-cl087 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl087-6590.5' similar to TR:Q9SSV4 Q9SSV4
MYO-INOSITOL-1-PHOSPHATE SYNTHASE.; mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

1 (bases 1 to 443)

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,

Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,

Beck, C., Wylie, T., Underwood, K., Stephens, M., Theising, B., Allen, M.,

Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,

McGinn, R., Ritter, E., Kohn, S., Shun, T., Jackson, Y., Cardenas, M.,

McCann, R., Waterson, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

When it has been determined, an EST from the other end of this

clone is listed in the 'Other ESTs on clone' field. This clone is

available through: Biogenetic Services, 801 32nd Ave. Brookings, SD

57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)

Seq primer: -40RP from Gibco

High quality sequence stop: 415.

FEATURES

source

1..443

/organism="Glycine max"

/mol_type="mRNA"

/culivar="Williams 82"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl087-6590"

/issue_type="Roots"

/lab_host="DH10B"

/clone_1lb="Gm-cl087"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI. The mRNA was prepared using polyactract mRNA system

from PROMEGA. The cDNA was prepared using the STRATAGENE

kit. Complementary DNA was synthesized from mRNA using a

primer consisting of a poly(dT) sequence with a XhoI

restriction site

(5'GAGAGAGAGAGAGAGAGAACTAGCTCTGACTTTTATTTTATTTT).

EcoRI adapters (5'OH-AATTCGACAGAG and 3'GCCGTGCTCp) were

ligated to the blunt-ended cDNA fragments followed by XhoI

digestion. The cDNA fragments were directionally cloned

into the EcoRI-XhoI restriction sites of the pBluescript

vector. The vector was previously dephosphorylated. The ligated cDNA fragments were transformed into DH10B by electroporation. Library is in LB medium with ampicillin and glycerol 8%. Average insert size: 800 bp."

ORIGIN

Query Match 25.2%; Score 443; DB 5; Length 443;
Best Local Similarity 100.0%; Pred. No. 8.2e-238; Indels 0; Gaps 0;
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 TGGCAATGCTTGTATGAAATGTTCTTTTCAATTAAGAGCCCTCAGAACACTTT 913
DB 1 TCCCAATGCTTGTATGAAATGTTCTTTTCAATTAAGAGCCCTCAGAACACTTT 60
QY 914 TGTACCAAGGCTGATGATCTTGGCCAGGAGAACACTTTGATGCTGAGATGACTT 973
DB 61 TGTACCAAGGCTGATGATCTTGGCCAGGAGAACACTTTGATGCTGAGATGACTT 120
QY 974 CAAGAGTGGTCAGACCAAAATGAAATCTGTGTTGTTGTTCTTGTGGGGCTGGTAT 1033
DB 121 CAAGAGTGGTCAGACCAAAATGAAATCTGTGTTGTTGTTCTTGTGGGGCTGGTAT 180
QY 1034 CAAGCCACATCTAATAGTCAAGTTCACCAATCTGGGAAACAATGATGATGATCTTTC 1093
DB 181 CAAGCCACATCTAATAGTCAAGTTCACCAATCTGGGAAACAATGATGATGATCTTTC 240
QY 1094 GGGTCACCAAACTTTCCGTTCCAGAGAAATCTTCCAGAGAACCTGTTGATGATGAT 1153
DB 241 GGGTCACCAAACTTTCCGTTCCAGAGAAATCTTCCAGAGAACCTGTTGATGATGAT 300
QY 1154 CAACAGCATGCGATCTCTATGAGCGCTGGTGAACATCCAGACATGTTGTTATTTAA 1213
DB 301 CAACAGCATGCGATCTCTATGAGCGCTGGTGAACATCCAGACATGTTGTTATTTAA 360
QY 1214 GATATGCTTTCATGATGAGGAGACAGCAAGAGCCATGATGATGATGATGATGAT 1273
DB 361 GATATGCTTTCATGATGAGGAGACAGCAAGAGCCATGATGATGATGATGATGAT 420
QY 1274 CATGGGTGAAAGAGACCAATTG 1296
DB 421 CATGGGTGAAAGAGACCAATTG 443

RESULT 13
BI347339 420 bp mRNA linear EST 30-OCT-2001
LOCUS F165B19 cDNA library from Forrest roots infected by F. solani
DEFINITION Glycine max cDNA clone F165B19, mRNA sequence.
ACCESSION BI347339
VERSION BI347339.1 GI:16519554
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE 1 (bases 1 to 420)
AUTHORS Iqbal,M.J. and Lightfoot,D.A.
TITLE Differentially expressed mRNA sequences identified by differential display of mRNA from soybean, Glycine max (L.) Merr. var. Forrest, roots in response to Fusarium solani f. sp. glycines inoculation
JOURNAL Unpublished (2001)
COMMENT Contact: Iqbal MJ and Lightfoot, DA
Center of Excellence in Soybean Research, Teaching and Outreach
Southern Illinois University at Carbondale
Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
USA
Tel: 618 453-3121
Fax: 618 453-7457
Email: mj1qbal@siu.edu, ga4082@siu.edu
Seg primer: M13 Forward.
FEATURES Location/Qualifiers

source

1..420
/organism="Glycine max"
/mol_type="mRNA"
/cultiivar="Forrest"
/db_xref="taxon:3847"
/clone="F165B19"
/dev_stage="14 days after inoculation with F. solani"
/clone_1ib="cDNA library from Forrest roots infected by F. solani"
/note="Organ: Root; Vector: pGEM; Forrest roots were inoculated with Fusarium solani f. sp. glycines and samples were collected after 14 days of inoculation. Total RNA was used for cDNA synthesis using SMART PCR cDNA synthesis kit (CLONTECH, CA) and cloned in pGEM vector. ESTs were sequenced using M13 universal forward or reverse primer."

ORIGIN

Query Match 22.7%; Score 400; DB 4; Length 420;
Best Local Similarity 100.0%; Pred. No. 1.6e-213; Indels 0; Gaps 0;
Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTTCTTATTCCTTTTGTATTCATTCATCTTAACTTTTGGAAAATTAATGTCA 60
DB 21 CTCTTCTTATTCCTTTTGTATTCATTCATCTTAACTTTTGGAAAATTAATGTCA 80
QY 61 TCGAATATTTAAGGTGAGTGTCTTAATGTGAATGACGAGACTGAGATTCAGTCG 120
DB 81 TCGAATATTTAAGGTGAGTGTCTTAATGTGAATGACGAGACTGAGATTCAGTCG 140
QY 121 TGTACAACTGCAAAACAACCACTTGTTCAGAGAACAGAAATGCACTATCAGTGA 180
DB 141 TGTACAACTGCAAAACAACCACTTGTTCAGAGAACAGAAATGCACTATCAGTGA 200
QY 181 TTGTCAAAACCAATCTGTCAATACGAATTTAAACCAATCATGTTCTTAATAG 240
DB 201 TTGTCAAAACCAATCTGTCAATACGAATTTAAACCAATCATGTTCTTAATAG 260
QY 241 GGGTATGCTTGTGGGTTGGGGTGAAACAACGCGCTCAACCTCAGCGTGGTGTAT 300
DB 261 GGGTATGCTTGTGGGTTGGGGTGAAACAACGCGCTCAACCTCAGCGTGGTGTAT 320
QY 301 CTACCGAGAGGGCAATTCATGGGCTACAAAGACAAATTCACAAAGCAATTAATT 360
DB 321 CTACCGAGAGGGCAATTCATGGGCTACAAAGACAAATTCACAAAGCAATTAATT 380
QY 361 GCTCCCTCACCCCAAGCTTACGATATCCGATGGGTCTT 400
DB 381 GCTCCCTCACCCCAAGCTTACGATATCCGATGGGTCTT 420

RESULT 14
BE609839 487 bp mRNA linear EST 13-UTL-2004
LOCUS sq45a10.y1 Gm-c1019 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1019-8371 5' similar to TR:064437 064437 MVO-INOSITOL PHOSPHATE SYNTHASE .; mRNA sequence.
ACCESSION BE609839
VERSION BE609839.1 GI:9900871
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 487)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvett,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurr,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE
JOURNAL
COMMENT

Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 When it has been determined, an EST from the other end of this
 clone is listed in the 'Other ESTs on clone' field. Trace
 considered overall poor quality. This clone is available through:
 Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone:
 800 423 4163; email: info@biogeneticservices.com)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 1.

FEATURES
source

1. 487
 Location/Qualifiers
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Williams"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1019-8371"
 /issue_type="Immature seed coats of greenhouse grown
 plants"
 /lab_host="DH10B (Gibco BRL)"
 /clone_lib="Gm-c1019"
 /note="Vector: pSPORT1 (Life Technologies); Site 1: Not I;
 Site 2: Sal I; This cDNA library was constructed from mRNA
 isolated from Immature seed coats (200-300 mg) of
 greenhouse grown plants. The library was prepared using
 the Life Technologies psuperscript cDNA library
 construction kit. Complementary DNA was synthesized from
 mRNA using a poly (dT) sequence with a Not I restriction
 site. Sal I linker adapters were ligated to the
 blunt-ended cDNA fragments followed by Not I digestion.
 The cDNA fragments were directionally cloned into the Not
 I-Sal I restriction site of the pSPORT vector. The
 ligated cDNA fragments were transformed into E.coli
 Electromax DH10B host cells (Gibco BRL). This library was
 constructed by Dr. Lila Vodka and Dr. Ann Khanna."

ORIGIN

Query Match 22.2%; Score 391; DB 2; Length 487;
 Best Local Similarity 99.8%; Pred. No. 2e-208; 1; Indels 0; Gaps 0;
 Matches 441; Conservative 0; Mismatches 1;

623 AGAGAGCGTGCACAAAGCTCATCAAGGCAAAAGCAAGCAAGTTCACAAATCAT 682
 1 AAGAGAGCGTGCACAAAGCTCATCAAGGCAAAAGCAAGCAAGTTCACAAATCAT 60
 683 CAAAGCATCAAGGCGTTTAAAGAGCCCAAAAGTGACAAAGGTGTGATCTGTGAC 742
 61 CAAAGCATCAAGGCGTTTAAAGAGCCCAAAAGTGACAAAGGTGTGATCTGTGAC 120
 743 TCCCAACAAAGAGGTACATTAATTTGGTGGGCTTTAATGACCAATGAGATCT 802
 121 TCCCAACAAAGAGGTACATTAATTTGGTGGGCTTTAATGACCAATGAGATCT 180
 803 CTGGGCTGTGTGACAGAAATAGAGCTGAGATTTCTCTTCCACCTTGTATGCCATTGC 862
 181 CTGGGCTGTGTGACAGAAATAGAGCTGAGATTTCTCTTCCACCTTGTATGCCATTGC 240
 863 TTGTGTTATGAAATGTTCTTTTCATTAATGAGACCCCTCAGAACACTTTGTACAG 922
 241 TTGTGTTATGAAATGTTCTTTTCATTAATGAGACCCCTCAGAACACTTTGTACAG 300
 923 GCTGATTAATCTTGCATGCGAGAAACCTTTGATTTGGTGGAGTACTTCAAGGTGG 982
 301 GCTGATTAATCTTGCATGCGAGAAACCTTTGATTTGGTGGAGTACTTCAAGGTGG 360
 983 TCAGACCAAAATGAATCTGTGTGTTGATTTCTTGTGGGGGCTGATCAAGCCAAC 1042
 |||||

Db 361 TCAGACCAAAATGAATCTGTGTGTTGATTTCTTGTGGGGCTGATCAAGCCAAC 420
 Oy 1043 ATCTATAGTCACTTACCAACAT 1064
 Db 421 ATCTATAGTCACTTACCAACAT 442

RESULT 15
 LOCUS AM472457 435 bp mRNA linear EST 14-JUL-2004
 DEFINITION s125e04.y1 Gm-c1029 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-c1029-2263 5' similar to SW:IN01 SPIPO P42803
 MYO-INOSITOL-1-PHOSPHATE SYNTHASE ;, mRNA sequence.
 ACCESSION AM472457 GI:7042563
 VERSION AM472457
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max

REFERENCE
AUTHORS

1 (bases 1 to 435)
 Shoemaker, R., Keim, P., Vodka, L., Erpelting, J., Corryell, V.,
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N.,
 Schurk, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.
 Public Soybean EST Project
 Unpublished (1999)

TITLE
JOURNAL
COMMENT

Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 When it has been determined, an EST from the other end of this
 clone is listed in the 'Other ESTs on clone' field. This clone is
 available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
 Insert Length: 1108 Std Error: 0.00.

FEATURES
source

1. 435
 Location/Qualifiers
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Williams"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1029-2263"
 /issue_type="very young cotyledons of greenhouse grown
 plants"
 /lab_host="DH10B"
 /clone_lib="Gm-c1029"
 /note="Vector: pSPORT1, Site 1: Not I; Site 2: Sal I; This
 cDNA library was constructed from mRNA isolated from very
 young cotyledons (20-50mg fresh weight) of greenhouse
 grown plants. The library was prepared using the Life
 Technologies psuperscript cDNA library construction kit.
 Complementary DNA was synthesized from mRNA using a
 poly(dT) sequence with a Not I restriction site. Sal I
 linker adapters were ligated to the blunt-ended cDNA
 fragments followed by Not I digestion. The cDNA fragments
 were directionally cloned into the Not I-Sal I restriction
 site of the pSPORT vector. The ligated cDNA fragments
 were transformed into E.coli Electromax DH10B host cells.
 This library was constructed by Dr. Paul Keim and Dr.
 Virginia Corryell."

ORIGIN

Query Match 21.8%; Score 383; DB 2; Length 435;
 Best Local Similarity 99.8%; Pred. No. 6.4e-204;
 Matches 433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
OY      698 GTTTAAGAGGACCAAGTGAACAAGTGGTTGTACTGTGAACTGCCAACACAGAGAG 757
      |||||||
Db      2 GTTTAAGAGGACCAAGTGAACAAGTGGTTGTACTGTGAACTGCCAACACAGAGAG 61
      |||||||
OY      758 GTACAGTAATTTGGTTGGGCTTATGACACCATGAGAAATCTTGGCTGTGTGA 817
      |||||||
Db      62 GTACAGTAATTTGGTTGGGCTTATGACACCATGAGAAATCTTGGCTGTGTGA 121
      |||||||
OY      818 CAGAAATGAGGCTGAGATTCTCCCTTGCACCTTGATGCAATTGCTTGTGTATGGA 877
      |||||||
Db      122 CAGAAATGAGGCTGAGATTCTCCCTTGCACCTTGATGCAATTGCTTGTATGGA 181
      |||||||
OY      878 TGTTCCTTCATTATGAGAGCCCTCAGAACATTTTGACAGGCTGATTGATCTTGC 937
      |||||||
Db      182 TGTTCCTTCATTATGAGAGCCCTCAGAACATTTTGACAGGCTGATTGATCTTGC 241
      |||||||
OY      938 CATCGGAGGAACTTTGATTGGTGAATGACCTTCAAGAGTGTGAGACCAAAATGAA 997
      |||||||
Db      242 CATCGGAGGAACTTTGATTGGTGAATGACCTTCAAGAGTGTGAGACCAAAATGAA 301
      |||||||
OY      998 ATCTGTGTGGTTGATTTCTGTGGGGCTGGTATCAAGCCAATCTATAGTCAGTTA 1057
      |||||||
Db      302 ATCTGTGTGGTTGATTTCTGTGGGGCTGGTATCAAGCCAATCTATAGTCAGTTA 361
      |||||||
OY      1058 CAACCATCTGGGAAACAATGATGTATGATCTTTGGCTCCACAACCTTCCGTTCCAA 1117
      |||||||
Db      362 CAACCATCTGGGAAACAATGATGTATGATCTTTGGCTCCACAACCTTCCGTTCCAA 421
      |||||||
OY      1118 GGAATCTCCAGA 1131
      |||||||
Db      422 GGAATCTCCAGA 435
      |||||||
```

Search completed: June 8, 2005, 14:33:44
Job time : 6002.12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:50:16 ; Search time 299.302 Seconds
(without alignments)
9621.891 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctcttctctctcttctgtc.....aatgtagtataattctgtc 1760

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
 - 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
 - 3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
 - 4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
 - 5: /cgn2_6/prodata/1/ina/6C_COMB.seq:*
 - 6: /cgn2_6/prodata/1/ina/6D_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	1.3	1931	3	US-09-118-442-10
2	23	1.3	1931	3	US-09-677-064-10
3	23	1.3	3546	3	US-09-118-442-14
4	23	1.3	3546	3	US-09-118-442-15
5	23	1.3	3546	3	US-09-677-064-14
6	23	1.3	3546	3	US-09-677-064-15
7	20	1.1	27223	4	US-09-949-016-13036
8	20	1.1	139049	4	US-09-949-016-17030
9	19	1.1	152132	4	US-09-175-928-9
10	19	1.1	152132	4	US-09-949-016-13845
11	19	1.1	152145	4	US-09-949-016-12371
12	19	1.1	374159	4	US-09-949-016-15868
13	19	1.1	1664976	4	US-08-916-421B-1
14	19	1.1	1664976	4	US-09-693-570-1
15	18	1.0	256	4	US-09-313-294A-2055
16	18	1.0	271	4	US-09-513-998C-9342
17	18	1.0	472	3	US-09-641-638-549
18	18	1.0	472	3	US-09-641-638-550
19	18	1.0	472	4	US-10-170-097-549
20	18	1.0	472	4	US-10-170-097-550
21	18	1.0	552	4	US-09-248-796A-13671
22	18	1.0	552	4	US-09-949-016-41180
23	18	1.0	601	4	US-09-949-016-141642
24	18	1.0	1001	3	US-09-641-638-278
25	18	1.0	1001	3	US-10-170-097-278
26	18	1.0	1062	3	US-09-668-680-10
27	18	1.0	1228	3	US-09-071-035-315

28	18	1.0	1326	4	US-09-134-000C-3096	Sequence 3096, App
29	18	1.0	1329	3	US-09-071-035-313	Sequence 313, App
30	18	1.0	2219	3	US-08-510-646B-17	Sequence 17, App1
31	18	1.0	2477	4	US-09-907-794A-169	Sequence 169, App
32	18	1.0	2477	4	US-09-905-125A-169	Sequence 169, App
33	18	1.0	2477	4	US-09-902-775A-169	Sequence 169, App
34	18	1.0	2477	4	US-09-906-700-169	Sequence 169, App
35	18	1.0	2477	4	US-09-903-603A-169	Sequence 169, App
36	18	1.0	2477	4	US-09-904-920A-169	Sequence 169, App
37	18	1.0	2477	4	US-09-909-064-169	Sequence 169, App
38	18	1.0	2477	4	US-09-905-381A-169	Sequence 169, App
39	18	1.0	2477	4	US-09-906-618-169	Sequence 169, App
40	18	1.0	4989	4	US-09-693-011-12	Sequence 12, App1
41	18	1.0	5083	4	US-09-693-011-11	Sequence 11, App1
42	18	1.0	6314	4	US-09-693-011-9	Sequence 9, App1
43	18	1.0	6408	4	US-09-693-011-10	Sequence 10, App1
44	18	1.0	16573	4	US-09-949-016-11764	Sequence 11764, A
45	18	1.0	17370	4	US-09-949-016-17331	Sequence 17331, A

ALIGNMENTS

```

RESULT 1
US-09-118-442-10
; Sequence 10, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Kang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FaestSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-10

Query Match          1.3%; Score 23; DB 3; Length 1931;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1158 AGCAATGCCATCCTCTATGAGCC 1180
Db      1203 AGCAATGCCATCCTCTATGAGCC 1225

RESULT 2
US-09-677-064-10
; Sequence 10, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064

```

```
/ CURRENT FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/055,446
/ PRIOR FILING DATE: 1997-08-11
/ PRIOR APPLICATION NUMBER: 60/055,526
/ PRIOR FILING DATE: 1997-08-08
/ PRIOR APPLICATION NUMBER: 60/053,944
/ PRIOR FILING DATE: 1997-07-28
/ PRIOR APPLICATION NUMBER: 09/118,442
/ PRIOR FILING DATE: 1998-07-17
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 10
/ LENGTH: 1931
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-677-064-10
```

```
Query Match 1.3%; Score 23; DB 3; Length 1931;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1158 AGCAATGCCATCCTCTATGAGCC 1180
DB 1203 AGCAATGCCATCCTCTATGAGCC 1225
```

```
RESULT 3
US-09-118-442-14
/ Sequence 14, Application US/09118442B
/ Patent No. 6197561
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ APPLICANT: Wang, Xun
/ APPLICANT: Bowen, Benjamin A.
/ TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
/ FILE REFERENCE: 0706
/ CURRENT APPLICATION NUMBER: US/09/118,442B
/ CURRENT FILING DATE: 1998-07-17
/ EARLIER APPLICATION NUMBER: 60/055,446
/ EARLIER FILING DATE: 1997-08-11
/ EARLIER APPLICATION NUMBER: 60/055,526
/ EARLIER FILING DATE: 1997-08-08
/ EARLIER APPLICATION NUMBER: 60/053,944
/ EARLIER FILING DATE: 1997-07-28
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 14
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-118-442-14
```

```
Query Match 1.3%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1158 AGCAATGCCATCCTCTATGAGCC 1180
DB 2842 AGCAATGCCATCCTCTATGAGCC 2864
```

```
RESULT 4
US-09-118-442-15
/ Sequence 15, Application US/09118442B
/ Patent No. 6197561
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ APPLICANT: Wang, Xun
```

```
/ APPLICANT: Bowen, Benjamin A.
/ TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
/ FILE REFERENCE: 0706
/ CURRENT APPLICATION NUMBER: US/09/118,442B
/ CURRENT FILING DATE: 1998-07-17
/ EARLIER APPLICATION NUMBER: 60/055,446
/ EARLIER FILING DATE: 1997-08-11
/ EARLIER APPLICATION NUMBER: 60/055,526
/ EARLIER FILING DATE: 1997-08-08
/ EARLIER APPLICATION NUMBER: 60/053,944
/ EARLIER FILING DATE: 1997-07-28
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 15
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-118-442-15
```

```
Query Match 1.3%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1158 AGCAATGCCATCCTCTATGAGCC 1180
DB 2842 AGCAATGCCATCCTCTATGAGCC 2864
```

```
RESULT 5
US-09-677-064-14
/ Sequence 14, Application US/09677064
/ Patent No. 6291224
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
/ APPLICANT: Wang, Hongyu
/ APPLICANT: Beach, Larry R.
/ TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
/ FILE REFERENCE: 0706D
/ CURRENT APPLICATION NUMBER: US/09/677,064
/ CURRENT FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/055,446
/ PRIOR FILING DATE: 1997-08-11
/ PRIOR APPLICATION NUMBER: 60/055,526
/ PRIOR FILING DATE: 1997-08-08
/ PRIOR APPLICATION NUMBER: 60/053,944
/ PRIOR FILING DATE: 1997-07-28
/ PRIOR APPLICATION NUMBER: 09/118,442
/ PRIOR FILING DATE: 1998-07-17
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 14
/ LENGTH: 3546
/ TYPE: DNA
/ ORGANISM: Zea mays
US-09-677-064-14
```

```
Query Match 1.3%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1158 AGCAATGCCATCCTCTATGAGCC 1180
DB 2842 AGCAATGCCATCCTCTATGAGCC 2864
```

```
RESULT 6
US-09-677-064-15
/ Sequence 15, Application US/09677064
/ Patent No. 6291224
/ GENERAL INFORMATION:
/ APPLICANT: Martino-Catt, Susan J.
```

APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.
TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
FILE REFERENCE: 0706D
CURRENT APPLICATION NUMBER: US/09/677,064
CURRENT FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/055,446
PRIOR FILING DATE: 1997-08-11
PRIOR APPLICATION NUMBER: 60/055,526
PRIOR FILING DATE: 1997-08-08
PRIOR APPLICATION NUMBER: 60/053,944
PRIOR FILING DATE: 1997-07-28
PRIOR APPLICATION NUMBER: 09/118,442
PRIOR FILING DATE: 1998-07-17
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 15
LENGTH: 3546
TYPE: DNA
ORGANISM: Zea mays
US-09-677-064-15

Query Match 1.3%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 AGCAATGCCATCCTCTATGAGCC 1180
DB 2842 AGCAATGCCATCCTCTATGAGCC 2864

RESULT 7
US-09-949-016-13036
Sequence 13036, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13036
LENGTH: 27223
TYPE: DNA
ORGANISM: Human
US-09-949-016-13036

Query Match 1.1%; Score 20; DB 4; Length 27223;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1529 AATGAGGCGTTGTGTGAT 1548
DB 3336 AATGAGGCGTTGTGTGAT 3355

RESULT 8
US-09-949-016-17030/c
Sequence 17030, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17030
LENGTH: 139049
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(139049)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17030

Query Match 1.1%; Score 20; DB 4; Length 139049;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 TTCTTAATTAGGGTATG 248
DB 30577 TTCTTAATTAGGGTATG 30558

RESULT 9
US-09-175-928-9/c
Sequence 9, Application US/09175928A
Patent No. 6312921
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Werberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Mi, Sha
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: 6006B.AJ172A
CURRENT APPLICATION NUMBER: US/09/175,928A
CURRENT FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 62
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 3153
TYPE: DNA
ORGANISM: Homo sapiens
US-09-175-928-9

Query Match 1.1%; Score 19; DB 3; Length 3153;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 TCACAGCATGCCATCCT 1171
DB 178 TCACAGCATGCCATCCT 160

RESULT 10
US-09-949-016-13845
Sequence 13845, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 13845
/ LENGTH: 152132
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-13845

Query Match          1.1%; Score 19; DB 4; Length 152132;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      860 TCCTTGTGTATGGAAT 878
Db      130119 TCCTTGTGTATGGAAT 130137

RESULT 11
US-09-949-016-12371
/ Sequence 12371, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 12371
/ LENGTH: 152145
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-12371

Query Match          1.1%; Score 19; DB 4; Length 152145;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      860 TCCTTGTGTATGGAAT 878
Db      130119 TCCTTGTGTATGGAAT 130137

RESULT 12
US-09-949-016-15868
/ Sequence 15868, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
```

```
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 15868
/ LENGTH: 374159
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-15868

Query Match          1.1%; Score 19; DB 4; Length 374159;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      755 GAGGTACAGTAATTTGGTT 773
Db      290127 GAGGTACAGTAATTTGGTT 290145

RESULT 13
US-08-916-421B-1
/ Sequence 1, Application US/08916421B
/ Patent No. 6503729
/ GENERAL INFORMATION:
/ APPLICANT: Bult et al.
/ TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
/ TITLE OF INVENTION: jannaschii
/ FILE REFERENCE: PB275
/ CURRENT APPLICATION NUMBER: US/08/916,421B
/ CURRENT FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: US 60/024,428
/ PRIOR FILING DATE: 1996-08-22
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 1
/ LENGTH: 1664976
/ TYPE: DNA
/ ORGANISM: Methanococcus jannaschii
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (28222)..(28222)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (28257)..(28258)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (84773)..(84773)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (84808)..(84808)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (84812)..(84812)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (98120)..(98120)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (98159)..(98159)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (98239)..(98239)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (98266)..(98266)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc_feature
/ LOCATION: (98343)..(98343)
/ OTHER INFORMATION: n equals a, t, c, or g
```



```

NAME/KEY: misc_feature
LOCATION: (103998) .. (103998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (148948) .. (148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385) .. (163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989) .. (191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191995) .. (191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (231980) .. (231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187) .. (234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220) .. (234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234814) .. (234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398) .. (309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418) .. (309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837) .. (312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312593) .. (312593)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226) .. (319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559157) .. (559157)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559241) .. (559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (600992) .. (600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708) .. (622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657081) .. (657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203) .. (657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435) .. (674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442) .. (682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713652) .. (713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684) .. (741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature

```

```

LOCATION: (779455) .. (779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779676) .. (779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539) .. (855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619) .. (871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830) .. (1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846) .. (1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881) .. (1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130988) .. (1130988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224) .. (1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473) .. (1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491) .. (1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091) .. (1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020) .. (1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912) .. (1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734) .. (1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998) .. (1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854) .. (1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match
Best Local Similarity 1.1%; Score 19; DB 4; Length 1664976;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 866 TGTATGGAATAATGTTCT 884
Db 326675 TGTATGGAATAATGTTCT 326693

RESULT 14
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Buit et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
; Patent No. 6797466
; TITLE OF INVENTION: jannaschi

```

```
/ FILE REFERENCE: PB275C1
/ CURRENT APPLICATION NUMBER: US/09/692,570
/ CURRENT FILING DATE: 2003-01-14
/ PRIOR APPLICATION NUMBER: US 60/024,428
/ PRIOR FILING DATE: 1996-08-22
/ PRIOR APPLICATION NUMBER: US 08/916,421
/ PRIOR FILING DATE: 1997-08-22
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1664976
/ TYPE: DNA
/ ORGANISM: Methanococcus jannaschii
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (28222)..(28222)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (28257)..(28258)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (84773)..(84773)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (84808)..(84808)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (84812)..(84812)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98120)..(98120)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98159)..(98159)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98239)..(98239)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98266)..(98266)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98343)..(98343)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (103998)..(103998)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (148948)..(148948)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (163385)..(163385)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (191989)..(191989)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (191995)..(191995)
/ OTHER INFORMATION: n equals a, t, c, or g
```

```
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (231980)..(231980)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (234187)..(234187)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (234720)..(234720)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (309398)..(309398)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (309418)..(309418)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (312837)..(312837)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (312993)..(312993)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (319226)..(319226)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (559167)..(559167)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (559241)..(559241)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (600992)..(600992)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (622708)..(622708)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (657081)..(657081)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (657203)..(657203)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (674435)..(674435)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (682442)..(682442)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (713652)..(713652)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
```

```

; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

```

```

Query Match 1.1%; Score 19; DB 4; Length 1664976;
Best Local Similarity 100.0%; Pred.No. 57;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 866 TGTATGAAATGTTCT 884
DB 326675 TGTATGAAATGTTCT 326693

```

```

RESULT 15
US-09-313-294A-2055
; Sequence 2055, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Ialigudi, Raghnath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 2055
; LENGTH: 256

```

```

; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700551970H1
US-09-313-294A-2055

```

```

Query Match 1.0%; Score 18; DB 4; Length 256;
Best Local Similarity 100.0%; Pred.No. 96;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 488 CAACATGAACCTGGCTGA 505
DB 99 CAACATGAACCTGGCTGA 116

```

```

Search completed: June 8, 2005, 14:43:08
Job time : 303.302 secs

```

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 11:26:41 ; Search time 1091.38 Seconds
(without alignments)
9956.820 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctcttattcctcttctgt.....aatgcagataatttctgt 1760

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 6046767 seqs, 3099530249 residues

Word size : 0

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1760	100.0	1760	14	US-10-025-003-1
2	1760	100.0	1760	19	US-10-718-952-1
3	1538	87.4	1989	18	US-10-424-599-70167
4	1533	87.1	1533	14	US-10-025-003-9
5	1533	87.1	1533	19	US-10-718-952-9
6	1482	84.2	1533	14	US-10-025-003-5
7	1482	84.2	1533	19	US-10-718-952-5

8	295	16.8	594	18	US-10-424-599-70166	Sequence 70166, A
9	276	15.7	661	18	US-10-424-599-116597	Sequence 116597, A
10	146	8.3	634	18	US-10-424-599-22663	Sequence 22663, A
11	146	8.3	1533	14	US-10-025-003-11	Sequence 11, Appl
12	146	8.3	1533	14	US-10-025-003-13	Sequence 13, Appl
13	146	8.3	1533	14	US-10-025-003-15	Sequence 15, Appl
14	146	8.3	1533	19	US-10-718-952-11	Sequence 11, Appl
15	146	8.3	1533	19	US-10-718-952-13	Sequence 13, Appl
16	146	8.3	1533	19	US-10-718-952-15	Sequence 15, Appl
17	146	8.3	2582	18	US-10-424-599-12021	Sequence 12021, A
18	121	6.9	844	18	US-10-424-599-70162	Sequence 70162, A
19	50	2.8	2018	18	US-10-424-599-12022	Sequence 12022, A
20	50	2.8	2582	18	US-10-424-599-12021	Sequence 12021, A
21	48	2.7	989	18	US-10-424-599-12019	Sequence 12019, A
22	47	2.7	355	18	US-10-424-599-38648	Sequence 38648, A
23	35	2.0	136	18	US-10-424-599-13935	Sequence 13935, A
24	30	1.7	553	18	US-10-424-599-78882	Sequence 78882, A
25	30	1.7	551	18	US-10-424-599-83330	Sequence 83330, A
26	29	1.6	377	11	US-09-733-6272A-2381	Sequence 2381, Ap
27	29	1.6	566	19	US-10-021-323-12655	Sequence 12655, A
28	29	1.6	618	19	US-10-021-323-17771	Sequence 17771, A
29	28	1.6	365	18	US-10-424-599-93575	Sequence 93575, A
30	28	1.6	262	18	US-10-424-599-94906	Sequence 94906, A
31	26	1.5	252	18	US-10-424-599-12017	Sequence 12017, A
32	26	1.5	1673	20	US-10-425-115-47180	Sequence 47180, A
33	26	1.5	1941	20	US-10-733-930-2091	Sequence 2091, Ap
34	24	1.4	35	14	US-10-025-003-3	Sequence 3, Appl
35	24	1.4	35	14	US-10-718-952-3	Sequence 3, Appl
36	24	1.4	1924	19	US-10-767-701-14714	Sequence 14714, A
37	23	1.3	260	18	US-10-424-599-36573	Sequence 36573, A
38	23	1.3	290	19	US-10-437-963-36182	Sequence 36182, A
39	23	1.3	345	11	US-09-732-6272A-3508	Sequence 3508, Ap
40	23	1.3	429	18	US-10-424-599-52735	Sequence 52735, A
41	23	1.3	457	17	US-10-341-961A-259	Sequence 259, App
42	23	1.3	557	19	US-10-021-323-12840	Sequence 12840, A
43	23	1.3	601	19	US-10-021-323-15424	Sequence 15424, A
44	23	1.3	605	19	US-10-021-323-15202	Sequence 15202, A
45	23	1.3	1998	18	US-10-425-114-7081	Sequence 7081, Ap

ALIGNMENTS

RESULT 1
US-10-025-003-1
; Sequence 1, Application US/10025003
; Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025, 003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/835, 751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 1
LENGTH: 1760
TYPE: DNA
ORGANISM: Glycine max
US-10-025-003-1
Query Match 100.0%; Score 1760; DB 14; Length 1760;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Oy 1 CTCTCTTATATCTTTTGTAATTCATTCATCTTAAATCTTTGTGAAAAATATGTTCA 60
Db 1 CTTCTTTTATCTTTTGTAATTCATTCATCTTAAATCTTTGTGAAAAATATGTTCA 60
Oy 61 TCGAATATTTTAAAGTTGAGTGTCTTAATGTGAATACCGAGACTGAGATTCAGTCCG 120
Db 61 TCGAATATTTTAAAGTTGAGTGTCTTAATGTGAATACCGAGACTGAGATTCAGTCCG 120
Oy 121 TGTACATCAGAAACCAACCGAATCTTTCACGAGAACGAAATGGACCTTACGTGSA 180
Db 121 TGTACATCAGAAACCAACCGAATCTTTCACGAGAACGAAATGGACCTTACGTGSA 180
Oy 181 TTGTCAAAACCAATCTGTCAATAAGAAATTTAAACCAATCCATGTTCTTAATAG 240
Db 181 TTGTCAAAACCAATCTGTCAATAAGAAATTTAAACCAATCCATGTTCTTAATAG 240
Oy 241 GGGTAATGTTGGGTGGGTGGGTAACAGGCTCAACCTCACCCTGAGTGTATTTG 300
Db 241 GGGTAATGTTGGGTGGGTGGGTAACAGGCTCAACCTCACCCTGAGTGTATTTG 300
Oy 301 CTAAACGAGAGGSCATTCATGGGCTAACAGAGAACAAAGTTCAAGACCAATTAATTG 360
Db 301 CTAAACGAGAGGSCATTCATGGGCTAACAGAGAACAAAGTTCAAGACCAATTAATTG 360
Oy 361 GCTCCCTCAACCAAGCTCAGCTATCCAGTGGGTCTTCCAGGAGAGAAATCTATG 420
Db 361 GCTCCCTCAACCAAGCTCAGCTATCCAGTGGGTCTTCCAGGAGAGAAATCTATG 420
Oy 421 CCCCAATCAAGAGCTGCTTCCAAATGTTAACCTGAGACAATGTTGGGGATGGG 480
Db 421 CCCCAATCAAGAGCTGCTTCCAAATGTTAACCTGAGACAATGTTGGGGATGGG 480
Oy 481 AATATAGCAACATGAACCTGGCTGATGCCATGGCCAGAGGCAAAAGTTTGAATGAT 540
Db 481 AATATAGCAACATGAACCTGGCTGATGCCATGGCCAGAGGCAAAAGTTTGAATGAT 540
Oy 541 TGCAGAGAGAGTTGAGGCTTATCATGATTCATGCTTCCATCCCGGAATCTATGAC 600
Db 541 TGCAGAGAGAGTTGAGGCTTATCATGATTCATGCTTCCATCCCGGAATCTATGAC 600
Oy 601 CGGATTTTCTGCTGCAACCAAGAGAGCTGCCAACAGTCAATCAAGGCAACAAAGC 660
Db 601 CGGATTTTCTGCTGCAACCAAGAGAGCTGCCAACAGTCAATCAAGGCAACAAAGC 660
Oy 661 AAGAGCAAGTTCAACAAATCATCAAGAGACATCAAGGCTTTAAGAGGCAACAAAGTG 720
Db 661 AAGAGCAAGTTCAACAAATCATCAAGAGACATCAAGGCTTTAAGAGGCAACAAAGTG 720
Oy 721 ACAAGGTGTTGTAAGTGTGATGCTGCCAACAGAGAGGTATGATTTGTTGTGGGC 780
Db 721 ACAAGGTGTTGTAAGTGTGATGCTGCCAACAGAGAGGTATGATTTGTTGTGGGC 780
Oy 781 TTAATGACACCATGAGAAATCTCTTGGCTGTGTGAGACAAATGAGGCTGATTTCTC 840
Db 781 TTAATGACACCATGAGAAATCTCTTGGCTGTGTGAGACAAATGAGGCTGATTTCTC 840
Oy 841 CTTCCACCTTGAATGCAATGCTGTGTGTAAGAAATGTTCTTCAATTAAGAGGC 900
Db 841 CTTCCACCTTGAATGCAATGCTGTGTGTAAGAAATGTTCTTCAATTAAGAGGC 900
Oy 901 CTCAAGAACATTTTGTACAGAGGCTGATGATCTTGCATCCGAGAACATTTGATTTG 960
Db 901 CTCAAGAACATTTTGTACAGAGGCTGATGATCTTGCATCCGAGAACATTTGATTTG 960
Oy 961 GTGAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTTGTTGTTCTTG 1020
Db 961 GTGAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTTGTTGTTCTTG 1020
Oy 1021 TGGGGGCTGTATCAAGCAACATCTATGTCAGTTTACCAATCTGGGAAACATGATG 1080
Db 1021 TGGGGGCTGTATCAAGCAACATCTATGTCAGTTTACCAATCTGGGAAACATGATG 1080
Oy 1081 GTATGAATCTTTCGGCTCCACAACTTTCGTTCCAGAGAAATCTCCAGAGCAAGTTG 1140
```

```
Db 1081 GTATGAATCTTTCGGCTCCACAACTTTCGTTCCAGAGAAATCTCCAGAGCAAGTTG 1140
Oy 1141 TTGATGATATGCTCAACGAAATGCATCCCTATGAGCCGTGTAACATCAGACCATG 1200
Db 1141 TTGATGATATGCTCAACGAAATGCATCCCTATGAGCCGTGTAACATCAGACCATG 1200
Oy 1201 TTGTTGTTAATTAAGTATGCTTACGTAGGAGGACAGCAAGAGCCATGATGATCA 1260
Db 1201 TTGTTGTTAATTAAGTATGCTTACGTAGGAGGACAGCAAGAGCCATGATGATCA 1260
Oy 1261 CTTAGAGATATTCATGAGGTTGAAAAGACACCATTTGTTTGCACAAACATGCCAGATT 1320
Db 1261 CTTAGAGATATTCATGAGGTTGAAAAGACACCATTTGTTTGCACAAACATGCCAGATT 1320
Oy 1321 CCTCTTATGCTGCTCTTATATCTTGAACCTTGCTCTTCTGAGCTCAGACTAGAA 1380
Db 1321 CCTCTTATGCTGCTCTTATATCTTGAACCTTGCTCTTCTGAGCTCAGACTAGAA 1380
Oy 1381 TCGATTTAAAGCTGAATAAGGGAATAATTCACCTATTCACCCAGTTGCTACATCC 1440
Db 1381 TCGATTTAAAGCTGAATAAGGGAATAATTCACCTATTCACCCAGTTGCTACATCC 1440
Oy 1441 TTAGCTTCAACCAAGCTCCTCTGTTCCACCGGTAACACAGTGTGATGATGATGTT 1500
Db 1441 TTAGCTTCAACCAAGCTCCTCTGTTCCACCGGTAACACAGTGTGATGATGATGTT 1500
Oy 1501 CAAAGCAAGCTGCAATGCTGAAAAACATTAAGAGGCTGTGTTGATTTGAGCCCAAG 1560
Db 1501 CAAAGCAAGCTGCAATGCTGAAAAACATTAAGAGGCTGTGTTGATTTGAGCCCAAG 1560
Oy 1561 AATAATGATTTCTGAGTACAGAGTGAAGCATGAGACCGAAGAAATATATGTTGGGATG 1620
Db 1561 AATAATGATTTCTGAGTACAGAGTGAAGCATGAGACCGAAGAAATATATGTTGGGATG 1620
Oy 1621 CTAAGCTGAATGTTTATGTTATATATATATATATATATATATATATATATATAT 1680
Db 1621 CTAAGCTGAATGTTTATGTTATATATATATATATATATATATATATATATATATAT 1680
Oy 1681 TGCATCAGCTTCAATTAATGCTTTAAGAGGAGGCAATCTGTTTACTGAGAACTGAATG 1740
Db 1681 TGCATCAGCTTCAATTAATGCTTTAAGAGGAGGCAATCTGTTTACTGAGAACTGAATG 1740
Oy 1741 AATGATATATATTTTGTGT 1760
Db 1741 AATGATATATATTTTGTGT 1760

RESULT 2
US-10-718-952-1
; Sequence 1, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-1
```

Query Match 100.0%; Score 1760; DB 19; Length 1760;
Best Local Similarity 100.0%; Pred. No.0;
Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTTATTCCTTTTGTATTTTCATTCATTCCTTATCTTTGTGAAAAATATGTTCA 60
DB 1 CTCCTTTATTCCTTTTGTATTTTCATTCATTCCTTATCTTTGTGAAAAATATGTTCA 60
QY 61 TCGAATTTTAAAGTTGAGTGTCTATATGTAAATACCGAGACTGAGATTCAGTCCG 120
DB 61 TCGAATTTTAAAGTTGAGTGTCTATATGTAAATACCGAGACTGAGATTCAGTCCG 120
QY 121 TGTACAACTACGAACCAACCGAATCTGTTACAGAAACAGAAATGGCACTTACAGTGA 180
DB 121 TGTACAACTACGAACCAACCGAATCTGTTACAGAAACAGAAATGGCACTTACAGTGA 180
QY 181 TTGTCAAAACCAATGTCTCAAAATAGCAATTTAAACCAATCATGTTCTTAAATTAG 240
DB 181 TTGTCAAAACCAATGTCTCAAAATAGCAATTTAAACCAATCATGTTCTTAAATTAG 240
QY 241 GGGTAATGCTTGTGGGTGGGTTGAAACAAAGGCTTCAACCTCACCGGTGTTATTG 300
DB 241 GGGTAATGCTTGTGGGTGGGTTGAAACAAAGGCTTCAACCTCACCGGTGTTATTG 300
QY 301 CTAAACGAGAGGCAATTTCAATGGGCTAACAAAGAACAAAGATTTCAACAAAGCAATTA 360
DB 301 CTAAACGAGAGGCAATTTCAATGGGCTAACAAAGAACAAAGATTTCAACAAAGCAATTA 360
QY 361 GCTCCCTCAACCAAGCCTCAGATATCCGAGTTGGGTCTTCCAGAGGAGAAATCTATG 420
DB 361 GCTCCCTCAACCAAGCCTCAGATATCCGAGTTGGGTCTTCCAGAGGAGAAATCTATG 420
QY 421 CCCCATTTCAAGAGCCTGCTTCCAAATGTTAACTTGAACAATTTGTTGGGGAATGG 480
DB 421 CCCCATTTCAAGAGCCTGCTTCCAAATGTTAACTTGAACAATTTGTTGGGGAATGG 480
QY 481 ATATCAAGCAATGAACTGGCTGATGCCATGGCCAGGGAAAGGTTTGAACATGATG 540
DB 481 ATATCAAGCAATGAACTGGCTGATGCCATGGCCAGGGAAAGGTTTGAACATGATG 540
QY 541 TCGAAGAGAGTTGAGGCTTATACATGAAATCCATGCTTCACTCCCGGAATCTATGACC 600
DB 541 TCGAAGAGAGTTGAGGCTTATACATGAAATCCATGCTTCACTCCCGGAATCTATGACC 600
QY 601 CGGATTTCAATGCTGCCAACCAAGAGAGGCTGCCAACCAAGCTCATCAAGGGCAAAAG 660
DB 601 CGGATTTCAATGCTGCCAACCAAGAGAGGCTGCCAACCAAGCTCATCAAGGGCAAAAG 660
QY 661 AAGAGCAAGTTCAAAATCATCAAAAGACATCAAGGCGTTTAAAGAAAGCCAAAGTGG 720
DB 661 AAGAGCAAGTTCAAAATCATCAAAAGACATCAAGGCGTTTAAAGAAAGCCAAAGTGG 720
QY 721 ACAAGTGGTTGACTGTGGAAGTCCCAACAGAGAGTACAGTAATTTGTTGGGACC 780
DB 721 ACAAGTGGTTGACTGTGGAAGTCCCAACAGAGAGTACAGTAATTTGTTGGGACC 780
QY 781 TTAATGACACCAATGAGAAATCTTGTGGCTGTGACAGAAATGAAGCTGAGATTTCTC 840
DB 781 TTAATGACACCAATGAGAAATCTTGTGGCTGTGACAGAAATGAAGCTGAGATTTCTC 840
QY 841 CTTCCACCTTGTATGTCATTTGTTGTATGAAAAATGTTCTTTCATTAATGAAAGCC 900
DB 841 CTTCCACCTTGTATGTCATTTGTTGTATGAAAAATGTTCTTTCATTAATGAAAGCC 900
QY 901 CTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTCCATCGGAGAGAACATTTGATG 960
DB 901 CTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTCCATCGGAGAGAACATTTGATG 960
QY 961 GTGAGAGATCACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGTTGATTTCTTG 1020
DB 961 GTGAGAGATCACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGTTGATTTCTTG 1020

QY 1021 TGGGGCTGGTATCAAGCCAAACATCTATAGTCAAGTTCACCATCTGGGAAACATGATG 1080
DB 1021 TGGGGCTGGTATCAAGCCAAACATCTATAGTCAAGTTCACCATCTGGGAAACATGATG 1080
QY 1081 GATGATCTTTTCGGCTCCACAACTTTCGTTCCAAAGAAATCTCCAAAGAGCAAGTTG 1140
DB 1081 GATGATCTTTTCGGCTCCACAACTTTCGTTCCAAAGAAATCTCCAAAGAGCAAGTTG 1140
QY 1141 TTGATGATATGCTCAACAGCATGCTATCTCTATAGAGCTGTGAACATCCAGCATG 1200
DB 1141 TTGATGATATGCTCAACAGCATGCTATCTCTATAGAGCTGTGAACATCCAGCATG 1200
QY 1201 TTGTGTTTAAAGTATGTCCTTATAGTGGGAGACGAAGAGGCAATGAGATGATCA 1260
DB 1201 TTGTGTTTAAAGTATGTCCTTATAGTGGGAGACGAAGAGGCAATGAGATGATCA 1260
QY 1261 CTTCAAGATATTTCAATGAGTGGTGAAGAGCACTTGTTCACACATGCGAGGATTT 1320
DB 1261 CTTCAAGATATTTCAATGAGTGGTGAAGAGCACTTGTTCACACATGCGAGGATTT 1320
QY 1321 CCTCTTACCTGCTCTATATCTTGAAGTGGTCTTCTTGTGAGCTCAGACTAGAA 1380
DB 1321 CCTCTTACCTGCTCTATATCTTGAAGTGGTCTTCTTGTGAGCTCAGACTAGAA 1380
QY 1381 TCGAGTTTAAAGTGAATGAGGGAATTCACATCTTCCACCAAGTTGCTAACATCC 1440
DB 1381 TCGAGTTTAAAGTGAATGAGGGAATTCACATCTTCCACCAAGTTGCTAACATCC 1440
QY 1441 TCACTACCTCAACCAAGGCTCTCTGTTCCACCGGGTACACAGAGTGAATGATGTTG 1500
DB 1441 TCACTACCTCAACCAAGGCTCTCTGTTCCACCGGGTACACAGAGTGAATGATGTTG 1500
QY 1501 CAAAGAGGCTGATGCTGGAAGAACATATAGAGGCTTGTGTTGATTTGCCCAAGAG 1560
DB 1501 CAAAGAGGCTGATGCTGGAAGAACATATAGAGGCTTGTGTTGATTTGCCCAAGAG 1560
QY 1561 ATTAACATGATTTCCAGATCAAGTGAAGATGAGACCGGAAGATATATAGTTGGGTAG 1620
DB 1561 ATTAACATGATTTCCAGATCAAGTGAAGATGAGACCGGAAGATATATAGTTGGGTAG 1620
QY 1621 CCTAGTGAATGTTTATGTTATATATATATGTTTGTCTTATATTTTCCAGATGTA 1680
DB 1621 CCTAGTGAATGTTTATGTTATATATATATGTTTGTCTTATATTTTCCAGATGTA 1680
QY 1681 TGCATCACTTCAATATCTTTAAGCGGGCATATCTGTTTACTAGAAACATGAATG 1740
DB 1681 TGCATCACTTCAATATCTTTAAGCGGGCATATCTGTTTACTAGAAACATGAATG 1740
QY 1741 AATGATGATTAATTTGTGT 1760
DB 1741 AATGATGATTAATTTGTGT 1760

RESULT 3
US-10-424-598-70167
; Sequence 70167, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5323) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 70167
; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT347_34372C.1
US-10-424-539-70167

Query Match 87.4%; Score 1538; DB 18; Length 1989;

Beet Local Similarity 99.8%; Pred. No. 0; Mismatches 2; Indels 1; Gaps 1;

Matches 1758; Conservative 0; Indels 1; Gaps 1;

```
1 CTCTCTTATTCCTTTGTAATTCATTCATTCCTTAATCTTTGTAATAATGTTCA 60
164 CTCTCTTATTCCTTTGTAATTCATTCATTCCTTAATCTTTGTAATAATGTTCA 223
61 TCGAAGATTTTAAAGTTGAGTGTCTTAATGTAAGTACCGAGACTGATTCAGTCCG 120
224 TCGAAGATTTTAAAGTTGAGTGTCTTAATGTAAGTACCGAGACTGATTCAGTCCG 283
121 TGTACAACTAGAAAACCAAGAACTGTTACAGAGAACGAAATGGCACTTACGTGA 180
284 TGTACAACTAGAAAACCAAGAACTGTTACAGAGAACGAAATGGCACTTACGTGA 343
181 TGTACAACTAGAAAACCAAGAACTGTTACAGAGAACGAAATGGCACTTACGTGA 240
344 TGTACAACTAGAAAACCAAGAACTGTTACAGAGAACGAAATGGCACTTACGTGA 403
241 GGGTAACTGTTGGGTT- GGGGTGAAACCAAGGCTCAACCTTCAACGGTGTAT 299
404 GGGTAACTGTTGGGTTGGGGGTGAAACCAAGGCTCAACCTTCAACGGTGTAT 463
300 GGTAAACCGAGAGGGAATTCATGAGGCTACAAAGAACAAATTCACAAAGCAATTA 359
464 GGTAAACCGAGAGGGAATTCATGAGGCTACAAAGAACAAATTCACAAAGCAATTA 523
360 GGGTCCCTCAACCAAGGCTCAAGTATCCAGTGGTGGTCTTCCAGGAGAGAAATCTAT 419
524 GGGTCCCTCAACCAAGGCTCAAGTATCCAGTGGTGGTCTTCCAGGAGAGAAATCTAT 583
420 GGGTCCCTCAACCAAGGCTCAAGTATCCAGTGGTGGTCTTCCAGGAGAGAAATCTAT 479
584 GGGTCCCTCAACCAAGGCTCAAGTATCCAGTGGTGGTCTTCCAGGAGAGAAATCTAT 643
480 GATATCAGAAATGAACTGAGTGTGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 539
644 GATATCAGAAATGAACTGAGTGTGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 703
540 TTGCAAGAACAGTTGAGGCTTATACATGAAATCATCTCTCCCGGAAATCTATGAC 599
704 TTGCAAGAACAGTTGAGGCTTATACATGAAATCATCTCTCCCGGAAATCTATGAC 763
600 CCGGATTTTATGCTGCTCCCAAGAGAGGCTGCTCAACAGTCTCAAGGAGCACAAG 659
764 CCGGATTTTATGCTGCTCCCAAGAGAGGCTGCTCAACAGTCTCAAGGAGCACAAG 823
660 CAAGAGCAAGTTCAACCAATCATCAAGAGATCAAGAGGCTTAAAGAGCCCAAGAGT 719
824 CAAGAGCAAGTTCAACCAATCATCAAGAGATCAAGAGGCTTAAAGAGCCCAAGAGT 883
720 GACAAAGTGTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 779
884 GACAAAGTGTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 943
780 GTTAAATGACACATGAGAAATCTCTTGGCTGCTGTAAGACAAAGATGAGGCTGATTT 839
944 GTTAAATGACACATGAGAAATCTCTTGGCTGCTGTAAGACAAAGATGAGGCTGATTT 1003
840 CCTTCCACCTTGAAGCCATGCTGCTGTAAGAAAGTCTTCAATTAATGTAATGTAATG 899
1004 CCTTCCACCTTGAAGCCATGCTGCTGTAAGAAAGTCTTCAATTAATGTAATGTAATG 1063
900 CCTCAGAACTTTTGTACAGAGGCTGATGATCTTCCATGCGAGAGCACTTGTAT 959
1064 CCTCAGAACTTTTGTACAGAGGCTGATGATCTTCCATGCGAGAGCACTTGTAT 1123
960 GGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGATGATTTCTT 1019
```

```
1124 GGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGATTTCTT 1183
1020 GTGGGGGCTGGTATCAAGCCCAATCTATGATGATTAACCAATCTGGGAAACAATGAT 1079
1184 GTGGGGGCTGGTATCAAGCCCAATCTATGATGATTAACCAATCTGGGAAACAATGAT 1243
1080 GGTATGAATCTTTGGGCTTCCAAACCTTCCGTTTCCAAAGAAATCTCCAAAGCAAGTT 1139
1244 GGTATGAATCTTTGGGCTTCCAAACCTTCCGTTTCCAAAGAAATCTCCAAAGCAAGTT 1303
1140 GTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1199
1304 GTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1363
1200 GTTGTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
1364 GTTGTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1423
1260 ACTTCAAGATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319
1424 ACTTCAAGATATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1483
1320 TCCCTCTAGCTGCTCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1379
1484 TCCCTCTAGCTGCTCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1543
1380 ATCAGATTTAAAGCTGAAATGAGGAAATTCATCCTATTCACCCAGTTGCTTACATC 1439
1544 ATCAGATTTAAAGCTGAAATGAGGAAATTCATCCTATTCACCCAGTTGCTTACATC 1603
1440 CTGAGCTTCTTCAACCAAGGCTCTCTGTTTCCACCGGTTACACAGTGTGATGATGATGAT 1499
1604 CTGAGCTTCTTCAACCAAGGCTCTCTGTTTCCACCGGTTACACAGTGTGATGATGATGAT 1663
1500 TCNAAGCAGGCTGCAATCTGGAACCAATTAATGAGGCTGTTGTTGATGATGATGATGATGAT 1559
1664 TCNAAGCAGGCTGCAATCTGGAACCAATTAATGAGGCTGTTGTTGATGATGATGATGATGAT 1723
1560 AATAACATGATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1619
1724 AATAACATGATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1783
1620 GCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
1784 GCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1843
1680 ATGATCAGCTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739
1844 ATGATCAGCTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1903
1740 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1760
1904 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1924
```

```
RESULT 4
US-10-025-003-9
; Sequence 9, Application US/10025003
; Publication No. US2003074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
```


QY	54	ATGTTCAATCGAAGATTTTTAAAGTTGAGTGCTCTAAATGGAATGACACCGAATCTGAAGTT	113
Db	1	ATGTTCAATCGAAGATTTTTAAAGTTGAGTGCTCTAAATGGAATGACACCGAATCTGAAGTT	60
QY	114	CAGTCGCGTGTACAACTACGAAACCAACCCAACTTGTTTCAGAGAAACAGAAATGGACCTAT	173
Db	61	CAGTCGCGTGTACAACTACGAAACCAACCCAACTTGTTTCAGAGAAACAGAAATGGACCTAT	120
QY	174	CAGTGAATGTCAAAACCCAAATCTGTCAAAATACGAATTTAAACCAACATCATGTTCCT	233
Db	121	CAGTGAATGTCAAAACCCAAATCTGTCAAAATACGAATTTAAACCAACATCATGTTCCT	180
QY	234	AAATTAAGGGGTAATGCTTTGTGGGTTGGGGTGGAAAACAACGGCTACACCTCAACGGTGGT	293
Db	181	AAATTAAGGGGTAATGCTTTGTGGGTTGGGGTGGAAAACAACGGCTACACCTCAACGGTGGT	240
QY	294	GTATATTGCTAAACGAGAGGGGCAATTTCAATGGGCTACAAAGGACAGAATTCACCAAGCCAA	353
Db	241	GTATATTGCTAAACGAGAGGGGCAATTTCAATGGGCTACAAAGGACAGAATTCACCAAGCCAA	300
QY	354	TACTTTGGCTCCACACCAAGCCTCACTATCCGAGTTGGGTCCTTTCAGAGGAGAGAA	413
Db	301	TACTTTGGCTCCACACCAAGCCTCACTATCCGAGTTGGGTCCTTTCAGAGGAGAGAA	360
QY	414	ATCTATGCCCCATTCAAGAGCGCTGCTTCCAAATGTTAAACCTGACGACATTTGTGTTGGG	473
Db	361	ATCTATGCCCCATTCAAGAGCGCTGCTTCCAAATGTTAAACCTGACGACATTTGTGTTGGG	420
QY	474	GGATGGGATATACAGCAATGAAACCTGGCTGATCGCATGGCCAGGGCAAGATGTTTGAC	533
Db	421	GGATGGGATATACAGCAATGAAACCTGGCTGATCGCATGGCCAGGGCAAGATGTTTGAC	480
QY	534	ATCGATTTGGCAAGAGCAGTTGAGGCTCTTACATGGAATTCATGCTTCCACTCCCGGAATC	593
Db	481	ATCGATTTGGCAAGAGCAGTTGAGGCTCTTACATGGAATTCATGCTTCCACTCCCGGAATC	540
QY	594	TATGACCCCGATTTCAATTGCTGCCAACCAAGAGAGCGGTGCCAACAACGTATCAAGGGC	653
Db	541	TATGACCCCGATTTCAATTGCTGCCAACCAAGAGAGCGGTGCCAACAACGTATCAAGGGC	600
QY	654	ACAAAGCAAGGCAAGTTCACAAATCATCAAAACATCAAGGGGTTTAAGAAAGCCACC	713
Db	601	ACAAAGCAAGGCAAGTTCACAAATCATCAAAACATCAAGGGGTTTAAGAAAGCCACC	660
QY	714	AAAGTGCAACAAGGTGTGTACTGTGACTGTGACCTCCACAACAGAGAGGTACAGTAAATTTGGTT	773
Db	661	AAAGTGCAACAAGGTGTGTACTGTGACTGTGACCTCCACAACAGAGAGGTACAGTAAATTTGGTT	720
QY	774	GTGGGCTTTAATGACACCATGAGAGAATCTTTGGCTGCTGTGACACAGAAATGAGGCTGAG	833
Db	721	GTGGGCTTTAATGACACCATGAGAGAATCTTTGGCTGCTGTGTGACACAGAAATGAGGCTGAG	780
QY	834	ATTTCTCTCTTCAACCTTGTAATGCCATTTGCTTGTGTATGAAAAATGTTTCCTTTATTAT	893
Db	781	ATTTCTCTCTTCAACCTTGTAATGCCATTTGCTTGTGTATGAAAAATGTTTCCTTTATTAT	840
QY	894	GGAAGCCTCAGAACACTTTTGTATCCAGGGCTGATTAATCTTTGCAATCGAGAGAACACT	953
Db	841	GGAAGCCTCAGAACACTTTTGTATCCAGGGCTGATTAATCTTTGCAATCGAGAGAACACT	900
QY	954	TGTATGTGTGAGATGACTTCAGAGAGTGCTCAGCAAAATGAATCTGTGTGTGTGAT	1013

```

RESULT 5
US-10-718-952-9
; Sequence 9, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCARIDES AND PHYIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: max
US-10-718-952-9

```

Query Match	87.1%	Score 1533	DB 19	Length 1533
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1533; Conservative	0	Mismatches	0	Indels 0
				Gaps 0

```
QY 54 ATGTCATCGAGAATTTTAAGGTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATT 113
DB 1 ATGTCATCGAGAATTTTAAGGTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATT 60
QY 114 CAGTCCTGTGTACACTACGAAACCAACCGAATTGTTTACGAGAACGAAATGGCACTAT 173
DB 61 CAGTCCTGTGTACACTACGAAACCAACCGAATTGTTTACGAGAACGAAATGGCACTAT 120
QY 174 CAGTGAATGTGAAACCCAAATCTGTCAATGSAATTTAAACCAACATTCATGTTCT 233
DB 121 CAGTGAATGTGAAACCCAAATCTGTCAATGSAATTTAAACCAACATTCATGTTCT 180
QY 234 AAATTAGGGGTATGATGTTGGGTTGGGGTGGAAACACGAGCTCAACCTTCACCGGTGT 293
DB 181 AAATTAGGGGTATGATGTTGGGTTGGGGTGGAAACACGAGCTCAACCTTCACCGGTGT 240
QY 294 GTTATTTGCTAACCCGAGGGGCAATTTGATGGCTACAAAGAACAAAGATTTCAACAGCAAT 353
DB 241 GTTATTTGCTAACCCGAGGGGCAATTTGATGGCTACAAAGAACAAAGATTTCAACAGCAAT 300
QY 354 TACTTTGGCTCCCTCAACCCAGGCTCAGCTATCCGAGTTGGGCTTCAGAGGAGAGAA 413
DB 301 TACTTTGGCTCCCTCAACCCAGGCTCAGCTATCCGAGTTGGGCTTCAGAGGAGAGAA 360
QY 414 ATCTATGCCCATTCAGAGGCTGCTTCCATGTTTAACTTCAGACGACATTTGTTGGG 473
DB 361 ATCTATGCCCATTCAGAGGCTGCTTCCATGTTTAACTTCAGACGACATTTGTTGGG 420
QY 474 GGATGGGATATACGACCAATGAACTGGCTGATGCCATGGCCAGGCAAAAGTGTTCAC 533
DB 421 GGATGGGATATACGACCAATGAACTGGCTGATGCCATGGCCAGGCAAAAGTGTTCAC 480
QY 534 ATCGATTTGCAGAGAGTTGAGGCTTACATGGAATCCATGCTTCCCTCCCGGAATC 593
DB 481 ATCGATTTGCAGAGAGTTGAGGCTTACATGGAATCCATGCTTCCCTCCCGGAATC 540
QY 594 TATGACCCCGAATTTCAATGCTGCAACCAAGAGAGGCTGCCAACACCTCATCAAGGAC 653
DB 541 TATGACCCCGAATTTCAATGCTGCAACCAAGAGAGGCTGCCAACACCTCATCAAGGAC 600
QY 654 ACAAAAGCAAGAGAGTTCAACAATCATCAAAAGACATCAAGGCTTTAAGAGACCAAC 713
DB 601 ACAAAAGCAAGAGAGTTCAACAATCATCAAAAGACATCAAGGCTTTAAGAGACCAAC 660
QY 714 AAAGTGAAGAAGTGTGTACTGTGAGACTGCCAACAGAGAGTACAGTAATTTGGTT 773
DB 661 AAAGTGAAGAAGTGTGTACTGTGAGACTGCCAACAGAGAGTACAGTAATTTGGTT 720
QY 774 GTGGGCTTTAATGACACATGAGAACTCTTGGCTGTGTGAGACGAATGAGGCTGAG 833
DB 721 GTGGGCTTTAATGACACATGAGAACTCTTGGCTGTGTGAGACGAATGAGGCTGAG 780
QY 834 ATTTTCTCTTCCACTTGTATGCAATGCTGTGTGTATGAAAATGTTCTTTCATTAAT 893
DB 781 ATTTTCTCTTCCACTTGTATGCAATGCTGTGTGTATGAAAATGTTCTTTCATTAAT 840
QY 894 GGAAGCCCTCAGAACCTTTTGTACCAAGGAGTGAATGATCTTGCCATCGCAGGAACAT 953
DB 841 GGAAGCCCTCAGAACCTTTTGTACCAAGGAGTGAATGATCTTGCCATCGCAGGAACAT 900
QY 954 TTGATTTGTGAGATGACTTCAAGAGTGTGACCAAAATGAAATCTGTGTTGTTGAT 1013
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACCAAAATGAAATCTGTGTTGTTGAT 960
QY 1014 TTCTTTGTGGGGCTGTGATCAAGCCACATCTAATGTGATTAACACCATCTGGGAAAC 1073
DB 961 TTCTTTGTGGGGCTGTGATCAAGCCACATCTAATGTGATTAACACCATCTGGGAAAC 1020
QY 1074 AATGATGATGATGAATTTTGGCTCCACAACTTTCCTGTTCAAGGAATCTTCCAAAGAC 1133
DB 1021 AATGATGATGATGAATTTTGGCTCCACAACTTTCCTGTTCAAGGAATCTTCCAAAGAC 1080
```

```
QY 1134 AACGTTGTGATGATANGTCAACAGCAATGCCATCTCTATATGAGCCTGTGTGAATCTCA 1193
DB 1081 AACGTTGTGATGATANGTCAACAGCAATGCCATCTCTATGAGCCTGTGTGAATCTCA 1140
QY 1194 GACCATGTTGTTGTTAATTAAGTATGCTGCTTAACGAGGGGACAGCAAGAGCCATGAT 1253
DB 1141 GACCATGTTGTTGTTAATTAAGTATGCTGCTTAACGAGGGGACAGCAAGAGCCATGAT 1200
QY 1254 GAGTACACTTCAGAGATATTTACATGGGTGAAAGAGACCAATGTTTTCACAACATATGC 1313
DB 1201 GAGTACACTTCAGAGATATTTACATGGGTGAAAGAGACCAATGTTTTCACAACATATGC 1260
QY 1314 GAGATTCCTCTTATGCTGCTCTATTAATCTTGAATGTTGTTCTTCTGAGCTCAAC 1373
DB 1261 GAGATTCCTCTTATGCTGCTCTATTAATCTTGAATGTTGTTCTTCTGAGCTCAAC 1320
QY 1374 ACTAATTCAGGTTTAAGCTGAATAAGGGAAATTCACCTCATTCACCCAGTTGCT 1433
DB 1321 ACTAATTCAGGTTTAAGCTGAATAAGGGAAATTCACCTCATTCACCCAGTTGCT 1380
QY 1434 ACCATCTCAGCTACCTCACCAAGGCTCTCTGTTCCACCGGGTACACAGTGTGAT 1493
DB 1381 ACCATCTCAGCTACCTCACCAAGGCTCTCTGTTCCACCGGGTACACAGTGTGAT 1440
QY 1494 GCATTGTCAAGACGCTGTGCAATGCTGGAATAATATATGAGGCTTGTGTGATTCGC 1553
DB 1441 GCATTGTCAAGACGCTGTGCAATGCTGGAATAATATATGAGGCTTGTGTGATTCGC 1500
QY 1554 CCAGGAATTAACATGATTCGAGTACAAAGTGA 1586
DB 1501 CCAGGAATTAACATGATTCGAGTACAAAGTGA 1533
```

RESULT 6

```
US-10-025-003-5
; Sequence 5, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Steitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-025-003-5
```

```
Query Match 84.2%; Score 1482; DB 14; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 54 ATGTCATCGAGAATTTTAAGGTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATT 113
DB 1 ATGTCATCGAGAATTTTAAGGTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATT 60
QY 114 CAGTCCTGTGTACACTACGAAACCAACCGAATTGTTTACGAGAACGAAATGGCACTAT 173
DB 61 CAGTCCTGTGTACACTACGAAACCAACCGAATTGTTTACGAGAACGAAATGGCACTAT 120
QY 174 CAGTGAATGTGAAACCCAAATCTGTCAATGSAATTTAAACCAACATTCATGTTCT 233
```

Db 121 CAGTGGATTGTCAAACCAAACTCTGTCAAATACGAATTTAAACCAACATCATGTTCC 180
Qy 234 AAATTAGGGGTAATGCTTGGGTTGGGGTGGGAAACAACGGCTCAACCTCACCSTGGT 233
Db 181 AAATTAGGGGTAATGCTTGGGTTGGGGTGGGAAACAACGGCTCAACCTCACCSTGGT 240
Qy 294 GTTATTGCTAACCGAGAGGGCATTTTCATGGGCTACAAAGGACAAAGATTCAACAGCAAT 353
Db 241 GTTATTGCTAACCGAGAGGGCATTTTCATGGGCTACAAAGGACAAAGATTCAACAGCAAT 300
Qy 354 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCAGTTGGGTCCTTCCAGGAGAGAA 413
Db 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCAGTTGGGTCCTTCCAGGAGAGAA 360
Qy 414 ATCTATGCCCCATTCAAGAGCTGCTTCCAAATGCTTAACTTCAACGACATTTGTTGGG 473
Db 361 ATCTATGCCCCATTCAAGAGCTGCTTCCAAATGCTTAACTTCAACGACATTTGTTGGG 420
Qy 474 GGATGGGATATCAGAACATGAACCTGGCTGATGGCCAGGGCAAGGTTTGAAC 533
Db 421 GGATGGGATATCAGAACATGAACCTGGCTGATGGCCAGGGCAAGGTTTGAAC 480
Qy 534 ATCGATTTCAGAAAGAGTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAAATC 593
Db 481 ATCGATTTCAGAAAGAGTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAAATC 540
Qy 594 TATGACCCCGGATTTCTTGTCTGCCAACCAAGAGAGCTGCAACACCTCATCAAGGGC 653
Db 541 TATGACCCCGGATTTCTTGTCTGCCAACCAAGAGAGCTGCAACACCTCATCAAGGGC 600
Qy 654 ACAAGCAAGAGCAAGTTCAACAATCATCAAAAGATCAAGGCGTTTAAAGAAAGCAC 713
Db 601 ACAAGCAAGAGCAAGTTCAACAATCATCAAAAGATCAAGGCGTTTAAAGAAAGCAC 660
Qy 714 AAAGTGAACAAGGTGTTGTAAGTGTGAGTCTGCCAACACAGAGATACAGTAATTTGGT 773
Db 661 AAAGTGAACAAGGTGTTGTAAGTGTGAGTCTGCCAACACAGAGATACAGTAATTTGGT 720
Qy 774 GTGGGCTTAAATGACACCATGGAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 833
Db 721 GTGGGCTTAAATGACACCATGGAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
Qy 834 ATTTCTCTCCACCTTGTATGCAATGCTGTGTTATGGAATAATGCTTTCATTAAAT 893
Db 781 ATTTCTCTCCACCTTGTATGCAATGCTGTGTTATGGAATAATGCTTTCATTAAAT 840
Qy 894 GGAAGCCCTCAGAAACCTTTTGTATCAGAGGCTGATTTGATCTTGCATCGGAGAAACCT 953
Db 841 GGAAGCCCTCAGAAACCTTTTGTATCAGAGGCTGATTTGATCTTGCATCGGAGAAACCT 900
Qy 954 TTGATTGCTGAGATGACTTCAAGAGTGGTCAAGCAAAATGAATCTGTTGTTGAT 1013
Db 901 TTGATTGCTGAGATGACTTCAAGAGTGGTCAAGCAAAATGAATCTGTTGTTGAT 960
Qy 1014 TTCTCTGTGGGGCTGTATCAAGCAACATATGTCAAGTAAACAATCGGGAAAC 1073
Db 961 TTCTCTGTGGGGCTGTATCAAGCAACATATGTCAAGTAAACAATCGGGAAAC 1020
Qy 1074 AATGATGATGATATCTTCCGCTCCCAAACTTTCCTTCCAGGAAATCTTCCAGAGC 1133
Db 1021 AATGATGATGATATCTTCCGCTCCCAAACTTTCCTTCCAGGAAATCTTCCAGAGC 1080
Qy 1134 AACGTTGATGATATGCTCAACAGCAATGCAATCTCTATGAGCTGCTGGAACATCCA 1193
Db 1081 AACGTTGATGATATGCTCAACAGCAATGCAATCTCTATGAGCTGCTGGAACATCCA 1140
Qy 1194 GACCATGTTGTTTATTAAGTATGCTTACGTAGGGGACAGCAAGAGCATGAT 1253
Db 1141 GACCATGTTGTTTATTAAGTATGCTTACGTAGGGGACAGCAATGAGCCATGAT 1200
Qy 1254 GAGTACACTTCAAGATATTCATGAGGTGAAAGAGCAACATGTTTGGACAACATGC 1313
Db 1201 GAGTACACTTCAAGATATTCATGAGGTGAAAGAGCAACATGTTTGGACAACATGC 1260

Qy 1314 GAGGATTCCTCTAGCTGCTCTCTATATCTTGAATCTTGCTTGTGAGCTCAGC 1373
Db 1261 GAGGATTCCTCTAGCTGCTCTCTATATCTTGAATCTTGCTTGTGAGCTCAGC 1320
Qy 1374 ACTAGAAATGAGTTTAAAGCTGAAATGAGGAAATTCACACTCATTCACCAAGTTGCT 1433
Db 1321 ACTAGAAATGAGTTTAAAGCTGAAATGAGGAAATTCACACTCATTCACCAAGTTGCT 1380
Qy 1434 ACCATCTCAGCTACTCACCAGAGCTCCTTGTGTTCCACCGGTAACCAAGTGTGAAT 1493
Db 1381 ACCATCTCAGCTACTCACCAGAGCTCCTTGTGTTCCACCGGTAACCAAGTGTGAAT 1440
Qy 1494 GCATTGTCAAAGAGCTGCAATGCTGGAATAATGAGGCTTGTGATTTGGCC 1553
Db 1441 GCATTGTCAAAGAGCTGCAATGCTGGAATAATGAGGCTTGTGATTTGGCC 1500
Qy 1554 CCAGAGAAATACATGATTTCTCGATGTAACAAGTGA 1586
Db 1501 CCAGAGAAATACATGATTTCTCGATGTAACAAGTGA 1533

RESULT 7
US-10-718-952-5
; Sequence 5, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-718-952-5

Query Match 84.2%; Score 1482; DB 19; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 54 ANGTTCATGAGAAATTTTAAAGTTGAGTGTCTTAAATGTAATACACCGAGCTGAGATT 113
Db 1 ANGTTCATGAGAAATTTTAAAGTTGAGTGTCTTAAATGTAATACACCGAGCTGAGATT 60
Qy 114 CAGTCCGTGTACACTACGAAACACCGAACTTTGTTACGAGAAACAGAAATGGCACTTAT 173
Db 61 CAGTCCGTGTACACTACGAAACACCGAACTTTGTTACGAGAAACAGAAATGGCACTTAT 120
Qy 174 CAGTGGATTGTCAAAACCCAAATCTGTCAAAATGCAATTTTAAACCAACATCCATGTTCC 233
Db 121 CAGTGGATTGTCAAAACCCAAATCTGTCAAAATGCAATTTTAAACCAACATCCATGTTCC 180
Qy 234 AAATTAGGGGTAATGCTTGGGTTGGGGTGGGAAACAACGGCTCAACCTCACCSTGGT 293
Db 181 AAATTAGGGGTAATGCTTGGGTTGGGGTGGGAAACAACGGCTCAACCTCACCSTGGT 240
Qy 294 GTTATTGCTAACCGAGAGGGCATTTTCATGGGCTACAAAGGACAAAGATTCAACAGCAAT 353
Db 241 GTTATTGCTAACCGAGAGGGCATTTTCATGGGCTACAAAGGACAAAGATTCAACAGCAAT 300
Qy 354 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCAGTTGGGTCCTTCCAGGAGAGAA 413

```
Db 301 TACTTGGCTCCCTCAACCCAGGCTCAGCTATCCGAGTGGGCTTCCAGGAGAGGAA 360
Qy 414 ATCTATGCCCATTCAGAGCCCTGCTTCCAAATGTTAAACCTGACACACTTGTGTTGGG 473
Db 361 ATCTATGCCCATTCAGAGCCCTGCTTCCAAATGTTAAACCTGACACACTTGTGTTGGG 420
Qy 474 GGATGGGATATGACAACTGAACTGGCTGATGCCATGGCCAGGCAAAAGGTTGTGAC 533
Db 421 GGATGGGATATGACAACTGAACTGGCTGATGCCATGGCCAGGCAAAAGGTTGTGAC 480
Qy 534 ATCGATTTTGCAGAGCAGTTGAGGCCCTTACATGGAATCCATGCTTCCACTCCCGGAATC 593
Db 481 ATCGATTTTGCAGAGCAGTTGAGGCCCTTACATGGAATCCATGCTTCCACTCCCGGAATC 540
Qy 594 TATGACCCCGGATTTCTATTGCTGCCAACAGAGAGCGCTCCAAACAAGTCTATCAAGGCG 653
Db 541 TATGACCCCGGATTTCTATTGCTGCCAACAGAGAGCGCTCCAAACAAGTCTATCAAGGCG 600
Qy 654 ACAGAGCAGAGCAGATTCACAAATCATCAAGACATCAAGGCGCTTTAAGAGAGCCACC 713
Db 601 ACAGAGCAGAGCAGATTCACAAATCATCAAGACATCAAGGCGCTTTAAGAGAGCCACC 660
Qy 714 AAAGTGGACAGGTGGTGTACTGTGAGCTGCCAACAGAGAGGTACAGTAATTTGGTT 773
Db 661 AAAGTGGACAGGTGGTGTACTGTGAGCTGCCAACAGAGAGGTACAGTAATTTGGTT 720
Qy 774 GTGGGCTTAAATGACACCAATGAGATCTTTGGCTGCTGAGACAGAAATAGAGGTGAG 833
Db 721 GTGGGCTTAAATGACACCAATGAGATCTTTGGCTGCTGAGACAGAAATAGAGGTGAG 780
Qy 834 ATTTCTCTTCCACCTTGTATGCGCATTTGTTGATGGAATAATGTTCTTCTTCAATTA 893
Db 781 ATTTCTCTTCCACCTTGTATGCGCATTTGTTGATGGAATAATGTTCTTCTTCAATTA 840
Qy 894 GGAAGCCTTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTCCATCGCGAGAACT 953
Db 841 GGAAGCCTTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTCCATCGCGAGAACT 900
Qy 954 TTGATTTGGGAGATGACCTTCAAGAGTGGTCAGACCAAAATGAAATCTGTTGTGTTAT 1013
Db 901 TTGATTTGGGAGATGACCTTCAAGAGTGGTCAGACCAAAATGAAATCTGTTGTGTTAT 960
Qy 1014 TTCTCTTGGGGGCTGGTATCAAGCCAAATCTATGTCAGTTACCAACCATCTGGGAAC 1073
Db 961 TTCTCTTGGGGGCTGGTATCAAGCCAAATCTATGTCAGTTACCAACCATCTGGGAAC 1020
Qy 1074 AATGATGATGAAATCTTCCGCTCCACAACTTTCCGTTCCAAAGGAAATCTTCCAGAGC 1133
Db 1021 AATGATGATGAAATCTTCCGCTCCACAACTTTCCGTTCCAAAGGAAATCTTCCAGAGC 1080
Qy 1134 AACGTTTGGATGATATGTCACACAGCATGCGATCTCTTATGAGCGCTGTGAAATCCA 1193
Db 1081 AACGTTTGGATGATATGTCACACAGCATGCGATCTCTTATGAGCGCTGTGAAATCCA 1140
Qy 1194 GACCATGTTGTTGTTAATGATATGCTTACGTAAGGAGGAGCAGCAAGAGAGCATGAT 1253
Db 1141 GACCATGTTGTTGTTAATGATATGCTTACGTAAGGAGGAGCAGCAATGAGCCATGAT 1200
Qy 1254 GAGTACACTTCAGAGATATTCATGGGTGAAAGAGACCATTTGTTTGCACAAACATGC 1313
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAGACCATTTGTTTGCACAAACATGC 1260
Qy 1314 GAGGATTCCTCTTACGCTCTCTATTAATCTTGAATTTGCTTCTTCTGAGCTCAGC 1373
Db 1261 GAGGATTCCTCTTACGCTCTCTATTAATCTTGAATTTGCTTCTTCTGAGCTCAGC 1320
Qy 1374 ACTAGAAATGAGTTTAAAGCTGAAATGAGGAGAAATTCACATCTTCCACCAAGTTGT 1433
Db 1321 ACTAGAAATGAGTTTAAAGCTGAAATGAGGAGAAATTCACATCTTCCACCAAGTTGT 1380
Qy 1434 ACCATCTCAGTACCTCAACAAAGCTCTCTGTTTCCACCGGATACACAGTGTGTAAT 1493
Db 1380 ACCATCTCAGTACCTCAACAAAGCTCTCTGTTTCCACCGGATACACAGTGTGTAAT 1340
```

```
Db 1381 ACCATCTCAGTACCTCAACAAAGCTCTCTGTTTCCACCGGATACACAGTGTGTAAT 1440
Qy 1494 GCATTGTCAAAGCAGCGTGCATGCTGAGAAAAATATATAGAGGCTTGTGTTGATGGCC 1553
Db 1441 GCATTGTCAAAGCAGCGTGCATGCTGAGAAAAATATATAGAGGCTTGTGTTGATGGCC 1500
Qy 1554 CCAGAGAAATTAACATGATTTCTCGAGTACAAAGTGA 1586
Db 1501 CCAGAGAAATTAACATGATTTCTCGAGTACAAAGTGA 1533

RESULT 8
US-10-424-599-70166
; Sequence 116597, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 70166
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34371C.1
US-10-424-599-70166

Query Match 16.8%; Score 295; DB 18; Length 594;
Best Local Similarity 100.0%; Pred. No. 1,1e-145; Mismatches 0; Gaps 0;
Matches 295; Conservative 0; Indels 0;

Qy 1466 GGTTCCACCGGGTACACCGAGTGTGATGATGTCATTTGCAAGAGCAGCGTCAATGCTGGAATA 1525
Db 300 GGTTCCACCGGGTACACCGAGTGTGATGATGTCATTTGCAAGAGCAGCGTCAATGCTGGAATA 359
Qy 1526 CATATAGAGGCTTGTGTTGATTTGCCCAAGAGATTAACATGATTTCTGAGTACAACTG 1585
Db 360 CATATAGAGGCTTGTGTTGATTTGCCCAAGAGATTAACATGATTTCTGAGTACAACTG 419
Qy 1586 AAGCATGGAGCCGAAGATTAATATGTTGGGTGACCTAGCTGAATGTTTATGTTAATA 1645
Db 420 AAGCATGGAGCCGAAGATTAATATGTTGGGTGACCTAGCTGAATGTTTATGTTAATA 479
Qy 1646 ATATGTTGCTTAATAATTTTGAAGTGAATGATGATGATGATGATGATGATGATGATGAT 1705
Db 480 ATATGTTGCTTAATAATTTTGAAGTGAATGATGATGATGATGATGATGATGATGATGAT 539
Qy 1706 AGCGGGGCAATATTTCTGTTTACAGGAACATGAATGAATGATTAATTTGTGT 1760
Db 540 AGCGGGGCAATATTTCTGTTTACAGGAACATGAATGAATGATTAATTTGTGT 594

RESULT 9
US-10-424-599-116597
; Sequence 116597, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
```

NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 116597
LENGTH: 661
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(661)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_76298C.1
US-10-424-599-116597

Query Match 15.7%; Score 276; DB 18; Length 661;
Best Local Similarity 100.0%; Pred.No.1.5e-135;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 818 CAGAAATGAGGCTGATGATTTCTCTCCACCTTGATGCGCATTTGTTATGAAAA 877
Db 4 CAGAAATGAGGCTGATGATTTCTCTCCACCTTGATGCGCATTTGTTATGAAAA 63
Qy 878 TGTTCCTTCATTATGAGAGCCCTCAGAACACTTTTGTACAGGGCTGATGATCTTGC 937
Db 64 TGTTCCTTCATTATGAGAGCCCTCAGAACACTTTTGTACAGGGCTGATGATCTTGC 123
Qy 938 CATCGGAGAACTTTGATTTGTTGATGATGATGATGATGATGATGATGATGATGAT 997
Db 124 CATCGGAGAACTTTGATTTGTTGATGATGATGATGATGATGATGATGATGATGAT 183
Qy 998 ATCTGTGTTGTTGATTTCTGTGGGGCTGGTATCAAGCCAACTATAGTCAGTTA 1057
Db 184 ATCTGTGTTGTTGATTTCTGTGGGGCTGGTATCAAGCCAACTATAGTCAGTTA 243
Qy 1058 CAACCATCTGGGAAACAATGATGATGATGATCTTTC 1093
Db 244 CAACCATCTGGGAAACAATGATGATGATGATCTTTC 279

RESULT 10

US-10-424-599-22663
Sequence 22663, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 22663
LENGTH: 634
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_120467C.1
US-10-424-599-22663

Query Match 8.3%; Score 146; DB 18; Length 634;
Best Local Similarity 100.0%; Pred.No.2.2e-66;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 870 ATGAAAAATGTTCTTCATTATGAGAGCCCTCAGAACACTTTTATACAGGGCTGATT 929
Db 116 ATGAAAAATGTTCTTCATTATGAGAGCCCTCAGAACACTTTTATACAGGGCTGATT 175
Qy 930 GATCTTGCCATGCGAGAAACACTTTGATGTTGATGATGATGATGATGATGATGATGAT 989
Db 176 GATCTTGCCATGCGAGAAACACTTTGATGTTGATGATGATGATGATGATGATGATGATGAT 235

Qy 990 AAAATGAATCTGTGTTGTTGATT 1015
Db 236 AAAATGAATCTGTGTTGTTGATT 261

RESULT 11

US-10-025-003-11
Sequence 11, Application US/10025003
Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 11
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-025-003-11

Query Match 8.3%; Score 146; DB 14; Length 1533;
Best Local Similarity 100.0%; Pred.No.2.3e-66;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 870 ATGAAAAATGTTCTTCATTATGAGAGCCCTCAGAACACTTTTATACAGGGCTGATT 929
Db 817 ATGAAAAATGTTCTTCATTATGAGAGCCCTCAGAACACTTTTATACAGGGCTGATT 876
Qy 930 GATCTTGCCATGCGAGAAACACTTTGATGTTGATGATGATGATGATGATGATGATGAT 989
Db 877 GATCTTGCCATGCGAGAAACACTTTGATGTTGATGATGATGATGATGATGATGATGATGAT 936
Qy 990 AAAATGAATCTGTGTTGTTGATT 1015
Db 937 AAAATGAATCTGTGTTGTTGATT 962

RESULT 12

US-10-025-003-13
Sequence 13, Application US/10025003
Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 13
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-025-003-13

Thu Jun 9 09:20:42 2005

us-10-718-952-1.011.rnpb

Page 11

Search completed: June 8, 2005, 19:34:52
Job time : 1093.38 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:47:55 ; Search time 92 Seconds
(without alignments)
2143.998 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 510
Sequence: 1 MFIEFKVCEVNVKTYETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq.16dec04.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	510	2	AAW79740 Soybean w
2	510	100.0	510	8	ADQ14439 Mutant so
3	510	100.0	510	8	ADQ14491 Wild type
4	510	100.0	510	8	ADG82000 Soybean m
5	510	100.0	510	8	ADG81994 Soybean m
6	409	80.2	510	2	AAW79741 Soybean m
7	409	80.2	510	8	ADQ14485 Mutant so
8	409	80.2	510	8	ADG81998 Soybean m
9	253	49.6	510	8	ADQ14503 Mutant so
10	253	49.6	510	8	ADQ14505 Wild type
11	253	49.6	510	8	ADG82004 Soybean m
12	253	49.6	510	8	ADG82006 Soybean m
13	236	46.3	510	8	ADQ14501 Mutant so
14	236	46.3	510	8	ADG82002 Soybean m
15	69	13.5	536	2	AAW79740 Soybean w
16	68	13.3	394	5	ABP51436 Human MDP
17	68	13.3	380	4	AAW79740 Soybean w
18	67	13.1	310	4	AAW79740 Soybean w
19	67	13.1	446	3	AAW79740 Soybean w
20	67	13.1	446	3	AAW79740 Soybean w
21	67	13.1	505	8	ADG73782 Myo-inosi
22	67	13.1	510	2	AAW79740 Soybean w
23	67	13.1	510	2	AAW79740 Soybean w
24	67	13.1	510	4	AAW79740 Soybean w
25	67	13.1	510	8	ADN73525 Thale cre

ALIGNMENTS

RESULT 1	AAW79740	standard; protein; 510 AA.	ADP43920 Rice L-my
ID	AAW79740	standard; protein; 510 AA.	AAG09861 Arabidops
XX	XX	AAW79740;	AAG32501 Arabidops
AC	XX	17-OCT-2003 (revised)	AAG09860 Arabidops
DT	DT	02-FEB-1999 (first entry)	AAG50574 Arabidops
XX	XX	soybean wild-type myo-inositol 1-phosphate synthase.	AAG50573 Arabidops
DE	XX	soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;	AAG52875 Arabidops
XX	XX	phytic acid.	AAG52874 Arabidops
KW	XX	Glycine max; line LR13.	
OS	XX	WO9845448-A1.	
PN	XX	15-OCT-1998.	
PD	XX	07-APR-1998; 98WO-US006822.	
XX	XX	08-APR-1997; 97US-00835751.	
PR	XX	(DUPO) DU PONT DE NEMOURS & CO E I.	
XX	XX	Hitz WD, Sebastian SA;	
PI	XX	WPI: 1998-568353/48.	
DR	XX	N-PSDB; AAW62440.	
DR	XX	Soybean plants containing altered myo-inositol 1-phosphate gene - useful	
XX	XX	for generating plants with altered levels of e.g. raffinose, stachyose,	
PT	XX	phytic acid, etc.	
PS	XX	Example 5; Page 45-47; 63pp; English.	
XX	XX	This is the amino acid sequence of soybean myo-inositol 1-phosphate	
CC	CC	synthase (MI 1-PS) deduced from the coding region of an isolated cDNA	
CC	CC	clone (see AAW62440). MI 1-PS is involved in glucose metabolism to phytic	
CC	CC	acid, raffinose and stachyose. A mutant MI 1-PS (see AAW79741) has been	
CC	CC	identified in soybean line LR33, a mutagenised line of low raffinose	
CC	CC	saccharide phenotype. Sequencing revealed a single base change mutation	
CC	CC	in the LR33 gene sequence that resulted in a K396N substitution in the	
CC	CC	mutant protein. The mutation results in a seed phenotype of very low	
CC	CC	raffinose saccharide sugars, very high sucrose and low phytic acid. The	
CC	CC	mutated nucleic acid is used to alter the raffinose saccharide, sucrose,	
CC	CC	phytic acid and inorganic phosphate content of soybean seeds, leading to	

CC useful soybean products, e.g. a seed phytic acid content of less than 17
CC ug/g, a seed content of raffinose and stachyose combined of less than
CC 14.5 ug/g, and a seed sucrose content greater than 200 ug/g. (Updated on
CC 17-OCT-2003 to standardise OS field)

XX Sequence 510 AA;

Query Match 100.0%; Score 510; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MFIEHFKECPNVKXTEIETISVYNYETTELHVNENRNGTYQWIVKPSVKYEFTKTNHVP 60
DB 1 MFIEHFKECPNVKXTEIETISVYNYETTELHVNENRNGTYQWIVKPSVKYEFTKTNHVP 60
OY 61 KLGVMLVGMGNNSTLTGVIANREGISMAWKDKIQOANYFGSLTQASAIRVGSFGQEE 120
DB 61 KLGVMLVGMGNNSTLTGVIANREGISMAWKDKIQOANYFGSLTQASAIRVGSFGQEE 120
OY 121 IYAPFKSLIPWNPDDIVFGGWDISNMMLADAMARAKYFDIDLQKQLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPWNPDDIVFGGWDISNMMLADAMARAKYFDIDLQKQLRPYMESMLPLPGI 180
OY 181 YDPDFIAANOERANNVIKGTQEOVOQI IKDIAFKKATKVDKVVVLTANTERYSNLV 240
DB 181 YDPDFIAANOERANNVIKGTQEOVOQI IKDIAFKKATKVDKVVVLTANTERYSNLV 240
OY 241 VGLNDTMENLLAAVDREABEISPSLTVAIACMVNPFPINSPONTFVPGILDLAIARNT 300
DB 241 VGLNDTMENLLAAVDREABEISPSLTVAIACMVNPFPINSPONTFVPGILDLAIARNT 300
OY 301 LIGGDDFKSGQTKMKSVLVDFLVAGAGIKPTISVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMKSVLVDFLVAGAGIKPTISVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
OY 361 NVVDDMVNSNALIYEPGHPDHVVVIKVPYVGSKRAMEYTSIIFMGKSTIVLHNTC 420
DB 361 NVVDDMVNSNALIYEPGHPDHVVVIKVPYVGSKRAMEYTSIIFMGKSTIVLHNTC 420
OY 421 EDSLILAAPIIIDVILVLAESTPIEFKANEKGKHSFHVATLTSLVTRAPLPPTPIVNV 480
DB 421 EDSLILAAPIIIDVILVLAESTPIEFKANEKGKHSFHVATLTSLVTRAPLPPTPIVNV 480
OY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 2

ADQ14499 standard, protein; 510 AA.

ADQ14499;

23-SEP-2004 (first entry)

Mutant soybean myo-inositol 1-phosphate synthase polypeptide #2.

CC Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
CC raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
CC mutant; mutant.

XX Glycine max.
OS Synthetic.

US2004128713-A1.

01-JUL-2004.

21-NOV-2003; 2003US-00718952.

08-APR-1997; 97US-00835751.
07-APR-1998; 98WO-US006822.

PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

XX WPI; 2004-53135/51.

DR N-PSDB; ADQ14498.

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

PS Example 8; SEQ ID NO 10; 48pp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.

XX Sequence 510 AA;

Query Match 100.0%; Score 510; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MFIEHFKECPNVKXTEIETISVYNYETTELHVNENRNGTYQWIVKPSVKYEFTKTNHVP 60
DB 1 MFIEHFKECPNVKXTEIETISVYNYETTELHVNENRNGTYQWIVKPSVKYEFTKTNHVP 60
OY 61 KLGVMLVGMGNNSTLTGVIANREGISMAWKDKIQOANYFGSLTQASAIRVGSFGQEE 120
DB 61 KLGVMLVGMGNNSTLTGVIANREGISMAWKDKIQOANYFGSLTQASAIRVGSFGQEE 120
OY 121 IYAPFKSLIPWNPDDIVFGGWDISNMMLADAMARAKYFDIDLQKQLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPWNPDDIVFGGWDISNMMLADAMARAKYFDIDLQKQLRPYMESMLPLPGI 180
OY 181 YDPDFIAANOERANNVIKGTQEOVOQI IKDIAFKKATKVDKVVVLTANTERYSNLV 240
DB 181 YDPDFIAANOERANNVIKGTQEOVOQI IKDIAFKKATKVDKVVVLTANTERYSNLV 240
OY 241 VGLNDTMENLLAAVDREABEISPSLTVAIACMVNPFPINSPONTFVPGILDLAIARNT 300
DB 241 VGLNDTMENLLAAVDREABEISPSLTVAIACMVNPFPINSPONTFVPGILDLAIARNT 300
OY 301 LIGGDDFKSGQTKMKSVLVDFLVAGAGIKPTISVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMKSVLVDFLVAGAGIKPTISVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
OY 361 NVVDDMVNSNALIYEPGHPDHVVVIKVPYVGSKRAMEYTSIIFMGKSTIVLHNTC 420
DB 361 NVVDDMVNSNALIYEPGHPDHVVVIKVPYVGSKRAMEYTSIIFMGKSTIVLHNTC 420
```

QY 421 EDSLLAAPITIDLVLAELSTRIEFKAENEGKFSFHPVATITLSYLTAKAPLVPGTPPVN 480
DB 421 EDSLLAAPITIDLVLAELSTRIEFKAENEGKFSFHPVATITLSYLTAKAPLVPGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510

RESULT 3
ADQ14491
ID ADQ14491 standard; protein; 510 AA.
AC ADQ14491;
DT 23-SEP-2004 (first entry)
DE Wild type soybean myo-inositol 1-phosphate synthase polypeptide #1.
XX Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KM raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.
XX Glycine max.
OS
XX US2004128713-A1.
PN 01-JUL-2004.
XX 21-NOV-2003; 2003US-00718952.
PF 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
PI MPI; 2004-533135/51.
DR N-PSDB; ADQ14490.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX Claim 3; SEQ ID NO 2; 48bp; English.

The invention relates to a nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate synthase having a decreased capacity for the synthesis of myo-inositol 1-phosphate. The invention also relates to a chimeric gene operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant comprising the chimeric gene, a method of making the soybean plant, a seed of the soybean plant, a soy protein product derived from the processing of soybean seeds, a method of making or producing a soy protein product and a method of using a soybean plant homozygous for at least one gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate. The nucleic acid fragment and methods are useful for producing plants with decreased raffinose, stachyose and phytic acid content and increased sucrose and inorganic phosphate content, leading to valuable and useful soybean products. This sequence represents a wild type soybean myo-inositol 1-phosphate synthase polypeptide of the invention.

Sequence 510 AA;

Query Match 100.0%; Score 510; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIENFKYECNPKYKTEIETISVYNYETTELVHNRNQTQYIMVPRKSVKTEFKTNINHP 60
DB 1 MEIENFKYECNPKYKTEIETISVYNYETTELVHNRNQTQYIMVPRKSVKTEFKTNINHP 60
QY 61 KLGVMLVGMGNGNSTLTGVIYANREGISMAITKDIQOANTFSGILTQASAIRVGSFOGEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIYANREGISMAITKDIQOANTFSGILTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLPWNPPDIFVFGGMDISNNMLADAMARAKVFDIDLOKQAPYMESMLPLPGI 180
DB 121 IYAPFKSLPWNPPDIFVFGGMDISNNMLADAMARAKVFDIDLOKQAPYMESMLPLPGI 180
QY 181 YDPDFIAANOEBRANNVIKGTKEQVOQI IKDIAFKKATKVDKVVMTANTERYSNLV 240
DB 181 YDPDFIAANOEBRANNVIKGTKEQVOQI IKDIAFKKATKVDKVVMTANTERYSNLV 240
QY 241 VGLNDTMENLLAAVDRNEAISPSTLVAIACWENVPIFGSPONTFVPGILDIAIANNT 300
DB 241 VGLNDTMENLLAAVDRNEAISPSTLVAIACWENVPIFGSPONTFVPGILDIAIANNT 300
QY 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSNHGNNDGMLSAPOFRSKISKS 360
DB 301 LIIGDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSNHGNNDGMLSAPOFRSKISKS 360
QY 361 NVVDDMVNSNAIILVEPGEHPHVVVYIKVPYVGDGSKRAMDEYTSIEIFMGKSTIYLAHNTC 420
DB 361 NVVDDMVNSNAIILVEPGEHPHVVVYIKVPYVGDGSKRAMDEYTSIEIFMGKSTIYLAHNTC 420
QY 421 EDSLLAAPITIDLVLAELSTRIEFKAENEGKFSFHPVATITLSYLTAKAPLVPGTPPVN 480
DB 421 EDSLLAAPITIDLVLAELSTRIEFKAENEGKFSFHPVATITLSYLTAKAPLVPGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510

RESULT 4
ADS82000
ID ADS82000 standard; protein; 510 AA.
AC ADS82000;
DT 18-NOV-2004 (first entry)
DE Soybean myo-inositol 1-phosphate synthase wild-type 3.
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
XX Glycine max; line 29004JPO1.
OS
XX US2003074685-A1.
PN 17-APR-2003.
PD 11-MAR-2002; 2002US-00025003.
PF 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
PI Hitz WD, Sebastian SA;
PI MPI; 2004-639957/62.
DR N-PSDB; ADS81999.
XX

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

PS Example 8; SEQ ID NO 10; 34pp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC oligosaccharides in soy plants (and other legumes) can lead to flautence
CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.

CC Sequence 510 AA;

Query Match 100.0%; Score 510; DB 8; Length 510;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;

Matches 510; Conservative 0; Indels 0;

QY 1 MFIENFKVCEPNVKKYETEISVYVYETTELVEHNRNGTYWIVPKSVKEFKTNHVP 60
DB 1 MFIENFKVCEPNVKKYETEISVYVYETTELVEHNRNGTYWIVPKSVKEFKTNHVP 60
QY 61 KLGVMLVWVGNGNGSTLTGGVIANREGISWATKDKIQOANFGSLTQASAIRVSGFQEE 120
DB 61 KLGVMLVWVGNGNGSTLTGGVIANREGISWATKDKIQOANFGSLTQASAIRVSGFQEE 120
QY 121 IYAPKSLLPWNPDPDIFVGGMDISNNMLADAMARAKYFDIDLQOLAPYVESMLPLGI 180
DB 121 IYAPKSLLPWNPDPDIFVGGMDISNNMLADAMARAKYFDIDLQOLAPYVESMLPLGI 180
QY 181 YDPDFIAANOEBERANNVKGTQEQVOOIIKDIKAFKATKVDKVVVMTANTERYSNLV 240
DB 181 YDPDFIAANOEBERANNVKGTQEQVOOIIKDIKAFKATKVDKVVVMTANTERYSNLV 240
QY 241 VGLNDTMENLLAADRNEAISPSTLYAIACWENVPTFINSPOPTFVPGILDIAIANNT 300
DB 241 VGLNDTMENLLAADRNEAISPSTLYAIACWENVPTFINSPOPTFVPGILDIAIANNT 300
QY 301 LIIGGDFESGOTKMSVLDPLVGAGIKPTSTVSVNHGNDNGMLSAPQFRSKYKISKS 360
DB 301 LIIGGDFESGOTKMSVLDPLVGAGIKPTSTVSVNHGNDNGMLSAPQFRSKYKISKS 360
QY 361 NVVDQMASNAILYEPGEHPDHVVIKVPYVGDSKRAMDEYTSIFMGKSTTYLHANTC 420
DB 361 NVVDQMASNAILYEPGEHPDHVVIKVPYVGDSKRAMDEYTSIFMGKSTTYLHANTC 420
QY 421 EDSLIIAAPTIIDLVLLAELSTRIEFKAENEGFHSFHPVATTILSYTKAPLVPPGTPVNV 480
DB 421 EDSLIIAAPTIIDLVLLAELSTRIEFKAENEGFHSFHPVATTILSYTKAPLVPPGTPVNV 480

DB 421 EDSLIIAAPTIIDLVLLAELSTRIEFKAENEGFHSFHPVATTILSYTKAPLVPPGTPVNV 480
QY 481 ALSKQRAMLENIMRACVGLAPENNMLEYK 510
DB 481 ALSKQRAMLENIMRACVGLAPENNMLEYK 510

RESULT 5
ADS81994
ID ADS81994 standard; protein; 510 AA.

AC ADS81994;

DT 18-NOV-2004 (first entry)

DE Soybean myo-inositol 1-phosphate synthase wild-type 1.

KW Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
KW raffinose; stachyose; sucrose; inorganic phosphate; flautence.

OS Glycine max; cultivar Wye.

PN US2003074685-A1.

PD 17-APR-2003.

PF 11-MAR-2002; 2002US-00025003.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.

PA (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PI Hitz WD, Sebastian SA;

DR WPI: 2004-639957/62.

DR N-PSDB; ADS81993.

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

PS Claim 3; SEQ ID NO 2; 34pp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC oligosaccharides in soy plants (and other legumes) can lead to flautence

CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 100.0%; Score 510; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MFIENFKECPNVKTETEISQSVNYVETTELVHNRNGTYQWIVPKSVKTEFKTNHVP 60
DB 1 MFIENFKECPNVKTETEISQSVNYVETTELVHNRNGTYQWIVPKSVKTEFKTNHVP 60
QY 61 KLGVMLVGMGNGNSTLTGVIYANREGISMAWKDKIQANFYGSLTQASAIRVGSFOGEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIYANREGISMAWKDKIQANFYGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPVNPDDIVFGGWDISNMMLADAMARAKYFDIDLOKQLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPVNPDDIVFGGWDISNMMLADAMARAKYFDIDLOKQLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEERANNVIKGTQEQVOQI IKDIFKFKATKVDKVVLTANTERYSNLV 240
DB 181 YDPDFIAANOEERANNVIKGTQEQVOQI IKDIFKFKATKVDKVVLTANTERYSNLV 240
QY 241 VGLNDTMENLLAAVDRNEAEISPSSTLYAIACMENVPPINGSPONTFVPGGLIDLAIAANT 300
DB 241 VGLNDTMENLLAAVDRNEAEISPSSTLYAIACMENVPPINGSPONTFVPGGLIDLAIAANT 300
QY 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSNHGNNDGMNLSAPQFRSKSISKS 360
DB 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSNHGNNDGMNLSAPQFRSKSISKS 360
QY 361 NVVDMDVNSNAIIVPEGHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVVDMDVNSNAIIVPEGHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLILAAPIIIDVLVLAELSTRIEFKANEKGKFSFHPVATILSYLTAKPLVPGETPVVN 480
DB 421 EDSLILAAPIIIDVLVLAELSTRIEFKANEKGKFSFHPVATILSYLTAKPLVPGETPVVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 6
AAW79741
ID AAW79741 standard; protein; 510 AA.

XX AC AAW79741;
XX DT 17-OCT-2003 (revised)
XX DT 02-FEB-1999 (first entry)
XX DE Soybean mutant myo-inositol 1-phosphate synthase.
XX KM Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
XX KM phytic acid.
XX OS Glycine max; line LR33.
XX PN MO9845448-A1.
XX PD 15-OCT-1998.
XX PF 07-APR-1998; 98MO-US006822.
XX PR 08-APR-1997; 97US-00835751.
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX PI Hiltz WD, Sebastian SA;

XX MPI: 1998-568353/48.
DR N-PSDB; AAV62443.
XX

PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
PT phytic acid, etc.

PS Example 5; Page 49-51; 63pp; English.

CC This is the amino acid sequence of a mutant soybean myo-inositol 1-
CC phosphate synthase (MI 1-PS) deduced from the coding region of an
CC isolated cDNA clone (see AAV62443). MI 1-PS is involved in glucose
CC metabolism to phytic acid, raffinose and stachyose. The MI 1-PS was
CC identified in soybean line LR33, a mutagenised line of low raffinose
CC saccharide phenotype. Sequencing revealed a single base change mutation
CC in the LR33 gene sequence that resulted in a K396N substitution in the
CC mutant protein compared to wild-type MI 1-PS (see AAW79740). The mutation
CC results in a seed phenotype of very low raffinose saccharide sugars, very
CC high sucrose and low phytic acid. The mutated nucleic acid is used to
CC alter the raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds, leading to useful soybean products,
CC e.g. a seed phytic acid content of less than 17 ug/g, a seed content of
CC raffinose and stachyose combined of less than 14.5 ug/g, and a seed
CC sucrose content greater than 200 ug/g. (Updated on 17-Oct-2003 to
CC standardise OS field)

XX Sequence 510 AA;

Query Match 80.2%; Score 409; DB 2; Length 510;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MFIENFKECPNVKTETEISQSVNYVETTELVHNRNGTYQWIVPKSVKTEFKTNHVP 60
DB 1 MFIENFKECPNVKTETEISQSVNYVETTELVHNRNGTYQWIVPKSVKTEFKTNHVP 60
QY 61 KLGVMLVGMGNGNSTLTGVIYANREGISMAWKDKIQANFYGSLTQASAIRVGSFOGEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIYANREGISMAWKDKIQANFYGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPVNPDDIVFGGWDISNMMLADAMARAKYFDIDLOKQLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPVNPDDIVFGGWDISNMMLADAMARAKYFDIDLOKQLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEERANNVIKGTQEQVOQI IKDIFKFKATKVDKVVLTANTERYSNLV 240
DB 181 YDPDFIAANOEERANNVIKGTQEQVOQI IKDIFKFKATKVDKVVLTANTERYSNLV 240
QY 241 VGLNDTMENLLAAVDRNEAEISPSSTLYAIACMENVPPINGSPONTFVPGGLIDLAIAANT 300
DB 241 VGLNDTMENLLAAVDRNEAEISPSSTLYAIACMENVPPINGSPONTFVPGGLIDLAIAANT 300
QY 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSNHGNNDGMNLSAPQFRSKSISKS 360
DB 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSNHGNNDGMNLSAPQFRSKSISKS 360
QY 361 NVVDMDVNSNAIIVPEGHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVVDMDVNSNAIIVPEGHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLILAAPIIIDVLVLAELSTRIEFKANEKGKFSFHPVATILSYLTAKPLVPGETPVVN 480
DB 421 EDSLILAAPIIIDVLVLAELSTRIEFKANEKGKFSFHPVATILSYLTAKPLVPGETPVVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 7
ADQ14495
ID ADQ14495 standard; protein; 510 AA.

```
XX ADQ14495;
XX
XX 23-SEP-2004 (first entry)
XX
XX Mutant soybean myo-inositol 1-phosphate synthase polypeptide #1.
XX
XX Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
XX raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
XX mutant; mutein.
XX
XX Glycine max.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 396 /note= "Wild type Lys substituted by Asn"
XX
XX US2004128713-A1.
XX
XX 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX 26-APR-1999; 99US-00299315.
XX 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX (GRAC/) GRACE D J.
XX (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
XX WPI; 2004-533135/51.
XX N-PSDB; ADQ14494.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX
XX Claim 9; SEQ ID NO 6; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents a mutant soybean myo-inositol
XX 1-phosphate synthase polypeptide of the invention.
XX
XX Sequence 510 AA:
XX
XX Query Match 80.2%; Score 409; DB 8; Length 510;
XX Best Local Similarity 99.8%; Pred. No. 0;
XX Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MFIEHFKECPNVKYTEETELIQSVYNYETTELVHENRNGTYQMIVKPKSVKTEFKTNHVP 60
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Dh 1 MFIEHFKECPNVKYTEETELIQSVYNYETTELVHENRNGTYQMIVKPKSVKTEFKTNHVP 60
Qy KLGVMVLVGWGGNNSTLTGVIANREGISWATKDKIQOANVFGSLTQASAIRVGSFOGEE 120
|||
Dh KLGVMVLVGWGGNNSTLTGVIANREGISWATKDKIQOANVFGSLTQASAIRVGSFOGEE 120
Qy IYAPFKSLLPVNVDDIVFGGWDISNNMLADAMARAKVPDIDLOKOLRPYMESMLPLPGI 180
121 IYAPFKSLLPVNVDDIVFGGWDISNNMLADAMARAKVPDIDLOKOLRPYMESMLPLPGI 180
Dh IYAPFKSLLPVNVDDIVFGGWDISNNMLADAMARAKVPDIDLOKOLRPYMESMLPLPGI 180
Qy YDPDFIAANOEBERANNVTKGTQEOVOQIITKDIKAFKEATVVDKVVVLTMTANTERYSNLV 240
181 YDPDFIAANOEBERANNVTKGTQEOVOQIITKDIKAFKEATVVDKVVVLTMTANTERYSNLV 240
Dh YDPDFIAANOEBERANNVTKGTQEOVOQIITKDIKAFKEATVVDKVVVLTMTANTERYSNLV 240
Qy VGLNDTMENLLAAVDRNEAEISPTLYAIACVMENVPFINGSPONTFVPGIDILAIARNT 300
241 VGLNDTMENLLAAVDRNEAEISPTLYAIACVMENVPFINGSPONTFVPGIDILAIARNT 300
Dh VGLNDTMENLLAAVDRNEAEISPTLYAIACVMENVPFINGSPONTFVPGIDILAIARNT 300
Qy LIGGDDFKSGQTKKKSVLVDFLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
301 LIGGDDFKSGQTKKKSVLVDFLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Dh LIGGDDFKSGQTKKKSVLVDFLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Qy NVVDDMNNSNAIIXEPGEHPDHVVVIVKVPYVGSXKRAMDEYSEIFMGKSTVLHNTC 420
361 NVVDDMNNSNAIIXEPGEHPDHVVVIVKVPYVGSXKRAMDEYSEIFMGKSTVLHNTC 420
Dh NVVDDMNNSNAIIXEPGEHPDHVVVIVKVPYVGSXKRAMDEYSEIFMGKSTVLHNTC 420
Qy EDSLILAPILIDVLVLAELSTRIEFKANEKGKFPHPVATILSYLTKAPLVPPGTPVNV 480
421 EDSLILAPILIDVLVLAELSTRIEFKANEKGKFPHPVATILSYLTKAPLVPPGTPVNV 480
Dh EDSLILAPILIDVLVLAELSTRIEFKANEKGKFPHPVATILSYLTKAPLVPPGTPVNV 480
Qy ALSKQRAMLENIMRACVGLAPENNMILLEYK 510
481 ALSKQRAMLENIMRACVGLAPENNMILLEYK 510
Dh ALSKQRAMLENIMRACVGLAPENNMILLEYK 510

RESULT 8
ADSB1998
ID ADSB1998 standard; protein; 510 AA.
XX
XX ADSB1998;
XX
XX 18-NOV-2004 (first entry)
XX
XX Soybean myo-inositol 1-phosphate synthase mutant #1.
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; mutein;
XX phytic acid; raffinose; stachyose; sucrose; inorganic phosphate;
XX flatulence; mutant.
XX
XX Glycine max; line LR33.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 396 /note= "Wild-type Lys substituted by Arg"
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
XX N-PSDB; ADSB1997.
XX
```

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

PS Claim 9; SEQ ID NO 6; 34bp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence represents a mutant myo-
CC inositol 1-phosphate synthase.

CC Sequence 510 AA;

Query Match 80.2%; Score 409; DB 8; Length 510;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFIEHFKCECPVVKTEIEISVYVYETTELVEHNRNGTYQIVPKSKYIEFKNIHP 60
DB 1 MFIEHFKCECPVVKTEIEISVYVYETTELVEHNRNGTYQIVPKSKYIEFKNIHP 60
QY 61 KLGVNLVGMGNGNSTLFGVIANREGISMATKDIQOANYGSLTQASAIRVGSFOGEE 120
DB 61 KLGVNLVGMGNGNSTLFGVIANREGISMATKDIQOANYGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPVNPDPDIVFGGWDISNMULADAMARAFVDIDLQOLRPYEMSLPLRGI 180
DB 121 IYAPFKSLIPVNPDPDIVFGGWDISNMULADAMARAFVDIDLQOLRPYEMSLPLRGI 180
QY 181 YDPPFIANOSBRANVVKGTQOEVOQIINDIKAFKATKVDKVVYMTANTERYSNLY 240
DB 181 YDPPFIANOSBRANVVKGTQOEVOQIINDIKAFKATKVDKVVYMTANTERYSNLY 240
QY 241 VGLNDTMENTLAAVDRNEAIEISPTLYAIACWENVPFINSPOFTFVGLDLIAIANT 300
DB 241 VGLNDTMENTLAAVDRNEAIEISPTLYAIACWENVPFINSPOFTFVGLDLIAIANT 300
QY 301 LIIGDDPKSGQTKMSVLVDPLVAGIKRPTISVYNHLGNDGMLSAPQFRSKEISKS 360
DB 301 LIIGDDPKSGQTKMSVLVDPLVAGIKRPTISVYNHLGNDGMLSAPQFRSKEISKS 360
QY 361 NVNDNMNSNAILVPERGHPDHVVVVKVPPYVGSKRAMDERTSITFMGKSTIYLAHTC 420
DB 361 NVNDNMNSNAILVPERGHPDHVVVVKVPPYVGSKRAMDERTSITFMGKSTIYLAHTC 420
QY 421 EDSLAAPIILDVLVLAELSTRIEFKANEKGFSHFVATILSYLTAPLPCTPVVN 480
DB 421 EDSLAAPIILDVLVLAELSTRIEFKANEKGFSHFVATILSYLTAPLPCTPVVN 480

DB 421 EDSLAAPIILDVLVLAELSTRIEFKANEKGFSHFVATILSYLTAPLPCTPVVN 480

QY 481 ALSKORAMLENTMRACVGLAPENNMLEYK 510

DB 481 ALSKORAMLENTMRACVGLAPENNMLEYK 510

RESULT 9

ADQ14503

ID ADQ14503 standard; protein; 510 AA.

AC ADQ14503;

DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #4.

KW Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;

KM raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;

KW mutant; mutein.

OS Glycine max.

PN US2004128713-A1.

PD 01-JUL-2004.

PF 21-NOV-2003; 2003US-00718952.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.

PR 26-APR-1999; 99US-0029315.

PR 11-MAR-2002; 2002US-00025003.

PA (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.

PA (STRE/) STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

PR MPI: 2004-533135/51.

PS N-PSDB; ADQ14502.

PS Example 8; SEQ ID NO 14; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.

CC Sequence 510 AA;

Query Match	49.6%;	Score 253;	DB 8;	Length 510;
Best Local Similarity	99.7%;	Pred.No.1.6e-244;		
Matches 353;	Conservative	0;	Mismatches 1;	Indels 0; Gaps 0
Qy	58	HVPKLGWLVGCGNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFO	117	
Dy	58	HVPKLGWLVGCGNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFO	117	
Qy	118	GEEIYAPFKSLPMVNPDDIVFGWDISNMNLADAMARAKVPDIDLQKQLRPYMESMLPL	177	
Dy	118	GEEIYAPFKSLPMVNPDDIVFGWDISNMNLADAMARAKVPDIDLQKQLRPYMESMLPL	177	
Qy	178	PGIYDPDPIAANOERANNVIGTQOEOVOQIIKQIKAFKATKYDKVYVLTANTERS	237	
Dy	178	PGIYDPDPIAANOERANNVIGTQOEOVOQIIKQIKAFKATKYDKVYVLTANTERS	237	
Qy	238	NLVVGLNDTMENTLAAVDRNEAEISPTLYIACVMENVPFNGSPONTFVGLDLAIA	297	
Dy	238	NLVVGLNDTMENTLAAVDRNEAEISPTLYIACVMENVPFNGSPONTFVGLDLAIA	297	
Qy	298	RNTLIGDDPFKSGQTKMKSVLDPFVAGIKPSTISVSYNHLGNDGMNISAPQTFRSKEI	357	
Dy	298	RNTLIGDDPFKSGQTKMKSVLDPFVAGIKPSTISVSYNHLGNDGMNISAPQTFRSKEI	357	
Qy	358	SKSNVVDVMVNSNALLIPEGEHPDHYVVIKYPYPYGDSCRANDEYTSSEIFMGK	411	
Dy	358	SKSNVVDVMVNSNALLIPEGEHPDHYVVIKYPYPYGDSCRANDEYTSSEIFMGK	411	
RESULT 10				
ID	ADQ14505	standard; protein, 510 AA.		
AC	ADQ14505;			
XX	23-SEP-2004	(first entry)		
XX	wild type soybean myo-inositol 1-phosphate synthase polypeptide #2.			
XX	Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;			
KW	raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.			
XX	Glycine max.			
XX	US2004128713-A1.			
PD	01-JUL-2004.			
XX	21-NOV-2003; 2003US-00718952.			
XX	08-APR-1997; 97US-00835751.			
PR	07-APR-1998; 98WO-US006822.			
PR	26-APR-1999; 98US-00299315.			
PR	11-MAR-2002; 2002US-00025003.			
XX	(HITZ// HITZ W D.			
PA	(SEBA// SEBASTIAN S A.			
PA	(GRAC// GRACE D J.			
PA	(STRE// STREIT L G.			
XX	Hitz WD, Sebastian SA, Grace DJ, Streit LG;			
XX	WPI; 2004-533135/51.			
DR	N-PSDB; ADQ14504.			
XX	New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,			
PT	useful for producing plants with decreased raffinose, stachyose, and			
PT	phytic acid and increased sucrose, leading to valuable and useful soybean			
XX	products.			
XX	Claim 3; SEQ ID NO 16; 48bp; English.			
CC	The invention relates to a nucleic acid fragment encoding a soybean myo-			

Query Match	49.6%	Score 253	DB 8	Length 510
Best Local Similarity	99.7%	Pred. No. 1.6e-244		
Matches 353	Conservative 0	Mismatches 1	Indels 0	Gaps 0
CC	inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate synthase having a decreased capacity for the synthesis of myo-inositol 1-phosphate. The invention also relates to a chimeric gene operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant comprising the chimeric gene, a method of making the soybean plant, a seed of the soybean plant, a soy protein product derived from the processing of soybean seeds, a method of making or producing a soy protein product and a method of using a soybean plant homozygous for at least one gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate. The nucleic acid fragment and methods are useful for producing plants with decreased raffinose, stachyose and phytic acid content and increased sucrose and inorganic phosphate content, leading to valuable and useful soybean products. This sequence represents a wild type soybean myo-inositol 1-phosphate synthase polypeptide of the invention.			
XX				
8Q	Sequence 510 AA:			
Query Match	49.6%	Score 253	DB 8	Length 510
Best Local Similarity	99.7%	Pred. No. 1.6e-244		
Matches 353	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	58	HYPKLGWLVGMGNGNSTLTGVIYANREGISMA TKDKIQANFSGSLTQASAIRVSGSQ	117	
DB	58	HYPKLGWLVGMGNGNSTLTGVIYANREGISMA TKDKIQANFSGSLTQASAIRVSGSQ	117	
QY	118	GEEIYAPFKSLPMVNPDDIVFGGMDISNMKLADAMARAKVFDIDLQQLRPYMESMPL	177	
DB	118	GEEIYAPFKSLPMVNPDDIVFGGMDISNMKLADAMARAKVFDIDLQQLRPYMESMPL	177	
QY	178	PGIYDPDFIAANOBERANNVIKGTKEQVOQIIKDIIKAFKATKVDKVVVIMLTANTERS	237	
DB	178	PGIYDPDFIAANOBERANNVIKGTKEQVOQIIKDIIKAFKATKVDKVVVIMLTANTERS	237	
QY	238	NLVVGLNDTMENLLAADVDRNAEISPSLTLYAACMEVVPINSGPONTFVPGLLDLAIA	297	
DB	238	NLVVGLNDTMENLLAADVDRNAEISPSLTLYAACMEVVPINSGPONTFVPGLLDLAIA	297	
QY	298	RNTLIIGDDPFSGGOTKMSVYVDFLVNAGIAPTSIVSNHLGNDGMNLAPQTPRSKEI	357	
DB	298	RNTLIIGDDPFSGGOTKMSVYVDFLVNAGIAPTSIVSNHLGNDGMNLAPQTPRSKEI	357	
QY	358	SKSNVVDVMVNSNALYEPGEHPDHVVVIKYVPYVGSKRAMEYTSIIFMGK	411	
DB	358	SKSNVVDVMVNSNALYEPGEHPDHVVVIKYVPYVGSKRAMEYTSIIFMGK	411	
Result 11				
ADSS82004	ID	ADSS82004	standard; protein; 510 AA.	
XX	XX	ADSS82004;		
XX	XX	18-NOV-2004 (first entry)		
DT	XX	Soybean myo-inositol 1-phosphate synthase wild-type 4.		
DE	XX			
XX	XX	Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;		
KM	XX	raffinose; stachyose; sucrose; inorganic phosphate; flatulence.		
XX	OS	Glycine max; line 2901BJP03.		
XX	PN	US2003074685-A1.		
XX	XX			
PD	XX	17-APR-2003.		
PF	XX	11-MAR-2002; 2002US-00025003.		
PR	XX	08-APR-1997; 97US-00835751.		
XX	PR	07-APR-1998; 98WO-US006822.		
XX	XX			

PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PI Hitz WD, Sebastian SA;
 XX
 XX WPI; 2004-639957/62.
 DR N-PSDB; ADS82003.
 XX

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
 PT phytic acid and inorganic phosphate content of soybean seeds.
 XX

PS Example 8; SEQ ID NO 14; 34pp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a
 CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 CC phosphate synthase having decreasing capacity for the synthesis for myo-
 CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
 CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 CC complement, subfragment or the complement of the subfragment, operably
 CC linked to, suitable regulatory sequences, where expression of the chimeric
 CC gene results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
 CC acid content of less than 17 micromol/g, a seed content of raffinose plus
 CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
 CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype
 CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homozygous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with
 CC heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence represents a wild-type myo-
 CC inositol 1-phosphate synthase.
 CC
 XX
 XX Sequence 510 AA;

Query Match 49.6%; Score 253; DB 8; Length 510;
 Best Local Similarity 99.7%; Pred. No. 1.6e-244;

Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 HVPKLGVMVGVGNGNGSTLTGVIANREGISWATKDKIQANFYGSLTQSAIRVSGSQ 117
 DB 58 HVPKLGVMVGVGNGNGSTLTGVIANREGISWATKDKIQANFYGSLTQSAIRVSGSQ 117
 QY 118 GEEIYAPFKSLPMPNPPDIYVGGWDISNMNLADAMARAKYVDILQKQLRPYMESMPL 177
 DB 118 GEEIYAPFKSLPMPNPPDIYVGGWDISNMNLADAMARAKYVDILQKQLRPYMESMPL 177
 QY 178 PGYIDPDIANQGERANNVIKGTQBOVOQIIKIKAFKATKDKVYVLTANTERS 237
 DB 178 PGYIDPDIANQGERANNVIKGTQBOVOQIIKIKAFKATKDKVYVLTANTERS 237
 QY 238 NLVVGNDTMENTLAAVDRNEAEISPTLYAIAQWENVPTFNGSPONTFVGLIDLAIA 297
 DB 238 NLVVGNDTMENTLAAVDRNEAEISPTLYAIAQWENVPTFNGSPONTFVGLIDLAIA 297
 QY 298 RNTLIGGDDFSGGTQKMSVLVDPLVAGIKRPTSTVSNHGLGNNGMLSAQOTRSKEI 357
 DB 298 RNTLIGGDDFSGGTQKMSVLVDPLVAGIKRPTSTVSNHGLGNNGMLSAQOTRSKEI 357
 QY 358 SKSNVVDMMVNSNALIPEGEHPDHVVVVKYVPYVGDSCRAMDYTSEIFMGCK 411
 DB 358 SKSNVVDMMVNSNALIPEGEHPDHVVVVKYVPYVGDSCRAMDYTSEIFMGCK 411

DB 358 SKSNVVDMMVNSNALIPEGEHPDHVVVVKYVPYVGDSCRAMDYTSEIFMGCK 411

RESULT 12

AD82006
 ID ADS82006 standard; protein; 510 AA.
 XX

AD82006;
 AC

18-NOV-2004 (first entry)
 DT

Soybean myo-inositol 1-phosphate synthase wild-type 2.
 DE

Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
 KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.

Glycine max; cultivar Wye.
 DE

US2003074685-A1.
 PN

17-APR-2003.
 PD

11-MAR-2002; 2002US-00025003.
 XX

08-APR-1997; 97US-00835751.
 XX

07-APR-1998; 98WO-US006822.
 PR

(HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.

Hitz WD, Sebastian SA;
 PI

WPI; 2004-639957/62.
 DR

N-PSDB; ADS82005.
 DR

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
 PT phytic acid and inorganic phosphate content of soybean seeds.
 PT

Claim 3; SEQ ID NO 16; 34pp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a
 CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 CC phosphate synthase having decreasing capacity for the synthesis for myo-
 CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
 CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 CC complement, subfragment or the complement of the subfragment, operably
 CC linked to suitable regulatory sequences, where expression of the chimeric
 CC gene results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
 CC acid content of less than 17 micromol/g, a seed content of raffinose plus
 CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
 CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype
 CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homozygous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with
 CC heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence represents a wild-type myo-
 CC inositol 1-phosphate synthase.
 CC
 XX
 XX Sequence 510 AA;

```
Query Match      49.6%; Score 253; DB 8; Length 510;
Best Local Similarity 99.7%; Pred. No. 1.6e-244;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 HVPKLGWLVWVGNGNGSTLTGCVIANKBGISWATKDKIQOANYGSLTQASAIRVGSFQ 117
DB 58 HVPKLGWLVWVGNGNGSTLTGCVIANKBGISWATKDKIQOANYGSLTQASAIRVGSFQ 117
QY 118 GBEIYAPKSLLPVNPDPDIYFGGMDISNMNLADMAKAPVDIDLOKQLRPYMSMLPL 177
DB 118 GBEIYAPKSLLPVNPDPDIYFGGMDISNMNLADMAKAPVDIDLOKQLRPYMSMLPL 177
QY 178 PGIVDPDFIAANOEBRANNVIKGTQEOVOQIIRKDKAFKATKYDKVVLMTANTERSYS 237
DB 178 PGIVDPDFIAANOEBRANNVIKGTQEOVOQIIRKDKAFKATKYDKVVLMTANTERSYS 237
QY 238 NLVVGINDTMENTLLAADVNRNEAISPSTLYAIACWENVPFINGSPQNTFVPGIDLAIA 297
DB 238 NLVVGINDTMENTLLAADVNRNEAISPSTLYAIACWENVPFINGSPQNTFVPGIDLAIA 297
QY 298 RNTLIGDDPKSGQTKMKSVLVDFLVGAGIKRPTISVSYNHLGNDGMMLSAPOFRSKKEI 357
DB 298 RNTLIGDDPKSGQTKMKSVLVDFLVGAGIKRPTISVSYNHLGNDGMMLSAPOFRSKKEI 357
QY 358 SKSNVVDVMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
DB 358 SKSNVVDVMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411

RESULT 13
ADQ14501
ID ADQ14501 standard; protein; 510 AA.
AC ADQ14501;
XX
DT 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #3.
XX
KW Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
KW mutant; mutein.
XX
OS Glycine max.
OS Synthetic.
XX
EN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
DR WP1; 2004-533135/51.
DR N-PSDB; ADQ14500.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
```

```
PS Claim 9; SEQ ID NO 12; 48bp; English.
XX
CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 510 AA:

Query Match      46.3%; Score 236; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.8e-227;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPDFIAANOEBRANNVIKGTQEOVOQIIRKDKAFKATKYDKVVLMTANTER 235
DB 176 PLPGIYDPDFIAANOEBRANNVIKGTQEOVOQIIRKDKAFKATKYDKVVLMTANTER 235
QY 236 YSNLVGINDTMENTLLAADVNRNEAISPSTLYAIACWENVPFINGSPQNTFVPGIDLA 295
DB 236 YSNLVGINDTMENTLLAADVNRNEAISPSTLYAIACWENVPFINGSPQNTFVPGIDLA 295
QY 296 IARNTLIGDDPKSGQTKMKSVLVDFLVGAGIKRPTISVSYNHLGNDGMMLSAPOFRSK 355
DB 296 IARNTLIGDDPKSGQTKMKSVLVDFLVGAGIKRPTISVSYNHLGNDGMMLSAPOFRSK 355
QY 356 EISKSNVVDVMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
DB 356 EISKSNVVDVMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411

RESULT 14
ADS82002
ID ADS82002 standard; protein; 510 AA.
XX
AC ADS82002;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase mutant #2.
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; mutein;
KW phytic acid; raffinose; stachyose; sucrose; inorganic phosphate;
KW flatulence; mutant.
XX
OS Glycine max; line 29010CP01.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 87 /note= "Wild-type Gly substituted by Asp"
XX
EN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
```

XX (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 XX
 XX
 PI Hitz WD, Sebastian SA;
 XX
 DR MPI: 2004-639957/62.
 DR N-PSDB; ADS82001.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

PS Claim 9; SEQ ID NO 12; 34pp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a mutant myo-inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 46.3%; Score 236; DB 8; Length 510;

Best Local Similarity 100.0%; Pred. No. 1.8e-227; Mismatches 0; Indels 0; Gaps 0;

DB 176 PLPGIYDPEPTAANQEBRANNVKKTQKQOVQIHKDIKAFKATKVDKVVVLTANTER 235
 176 PLPGIYDPEPTAANQEBRANNVKKTQKQOVQIHKDIKAFKATKVDKVVVLTANTER 235
 QY 236 YSNLVVGLNDTMENTLAAVDRENEAISPSTLYAIAQWENVPFINGSPQNTFVPGIDLA 295
 DB 236 YSNLVVGLNDTMENTLAAVDRENEAISPSTLYAIAQWENVPFINGSPQNTFVPGIDLA 295
 QY 296 IARNTLIGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYHNLGNNDGMNLSAPQTFRSK 355
 DB 296 IARNTLIGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYHNLGNNDGMNLSAPQTFRSK 355
 QY 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVVKYVPYVGDGSKRAMDEYTSEIFMGSK 411
 DB 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVVKYVPYVGDGSKRAMDEYTSEIFMGSK 411

RESULT 15
 AAY24477
 ID AAY24477 standard; protein; 536 AA.
 XX

AC AAY24477;

XX 24-SEP-1999 (first entry)

XX Nicotiana paniculata INPS protein.

XX Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;

XX water stress; resistance.

XX Nicotiana paniculata.

XX Key Location/Qualifiers

FT Misc-difference 511

FT //label= unknown

PN JP1187879-A.

XX 13-JUL-1999.

XX 26-DEC-1997; 97JP-00359773.

XX 26-DEC-1997; 97JP-00359773.

XX (N1SB) JAPAN TOBACCO INC.

XX MPI: 1999-451546/38.

DR N-PSDB; AAX90402.

XX New INPS gene derived from Nicotiana genus plant - useful for conferring resistance to water stress to plants.

XX Claim 2; Page 6-8; 8pp; Japanese.

XX The present sequence is the Nicotiana paniculata inositol monophosphate synthase (INPS), designated NpINPS1. INPS can be used to confer water stress resistance to a plant

XX Sequence 536 AA;

Query Match 13.5%; Score 69; DB 2; Length 536;

Best Local Similarity 100.0%; Pred. No. 6.2e-60; Mismatches 0; Indels 0; Gaps 0;

QY 299 NTLIGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYHNLGNNDGMNLSAPQTFRSKIS 358

DB 299 NTLIGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYHNLGNNDGMNLSAPQTFRSKIS 358

QY 359 KSNVVDMDV 367

DB 359 KSNVVDMDV 367

Search completed: June 7, 2005, 17:06:31
 Job time : 103 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:58:47 ; Search time 24.5 Seconds
(without alignments)
2002.880 Million cell updates/sec

Title: US-10-718-952-2

Sequence: 1 MFIENKVCSPNVKTYTETI.....NIMRACVGLAPENMILEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR 79: *
2: pir1: *
3: pir2: *
4: pir3: *
5: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	13.1	507	2	S52648 inositol-3-phospha
2	67	13.1	509	2	T08436 inositol-3-phospha
3	67	13.1	510	2	T04399 inositol-3-phospha
4	67	13.1	510	2	T50021 inositol-3-phospha
5	67	13.1	510	2	S60302 inositol-3-phospha
6	67	13.1	510	2	D84610 probable myo-inos
7	67	13.1	511	2	T05017 inositol-3-phospha
8	67	13.1	512	2	T12438 inositol-3-phospha
9	50	9.8	510	2	T01647 inositol-3-phospha
10	48	9.4	511	2	T10964 inositol-3-phospha
11	15	2.9	430	2	T46317 inositol-3-phospha
12	13	2.5	525	2	T18569 inositol-3-phospha
13	13	2.5	555	2	A30902 inositol-3-phospha
14	12	2.4	520	1	B45452 inositol-3-phospha
15	15	1.6	210	2	S25544 minor outer capsid
16	16	1.6	210	2	S25544 nonstructural prot
17	8	1.6	378	2	P95844 probable dihydrol
18	8	1.6	429	2	E75107 thiamin biosynthes
19	8	1.6	429	2	AB2340 hypochetrical prote
20	7	1.4	42	2	165746 tropomyosin - huma
21	7	1.4	96	1	MMVZP3 F3 protein - fowlp
22	7	1.4	97	2	Q00527 capsid assembly pr
23	7	1.4	105	2	B82149 conserved hypochet
24	7	1.4	125	2	AB1895 hypochetrical prote
25	7	1.4	147	2	T02273 hypochetrical prote
26	7	1.4	157	2	S77352 hypochetrical prote
27	7	1.4	178	2	S44910 ZK686.1 protein -
28	7	1.4	191	2	AH0431 probable exported
29	7	1.4	197	2	T15106 hypochetrical prote

30	7	1.4	201	2	AH1812 hypochetrical prote
31	7	1.4	209	2	A73313 probable v-type AT
32	7	1.4	212	2	T36864 50S ribosomal prot
33	7	1.4	231	2	B83111 hypochetrical prote
34	7	1.4	236	2	T33333 Machado-Joseph dis
35	7	1.4	280	2	T47572 rRNA methylase (im
36	7	1.4	282	2	AE2618 rRNA methylase (AP
37	7	1.4	290	2	D97400 ribose ABC transpo
38	7	1.4	294	2	C82497 hypochetrical prote
39	7	1.4	294	2	AE2457 rhamnosyl transfer
40	7	1.4	299	2	B84262 hypochetrical prote
41	7	1.4	300	2	S36430 tropomyosin - huma
42	7	1.4	308	2	T06796 probable short cha
43	7	1.4	309	2	C83017 hypochetrical prote
44	7	1.4	309	2	T02613 hydroxymethylgluta
45	7	1.4	311	2	D64209

ALIGNMENTS

RESULT 1

S52648 inositol-3-phosphate synthase (EC 5.5.1.4) - Citrus paradisi

C/Species: Citrus paradisi

C/Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004

C/Accession: S52648

R/Holland, D. submitted to the EMBL Data Library, April 1994

A/Reference number: S52648

A/Accession: S52648

A/Molecule type: DNA

A/Residues: 1-507 <HUS>

A/Cross-references: UNIPROT:P42802; GB:Z32632; NID:G602564; PIDN:CAA83565.1; PID:G602565

C/Genetics:

A/Gene: INO1

C/Superfamily: myo-inositol-1-phosphate synthase

C/Keywords: Intramolecular lyase; isomerase

Query Match 13.1%; Score 67; DB 2; Length 507;

Best Local Similarity 100.0%; Pred. No. 1.6e-60;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTRKMSVLVDPLVGAGIKPTYSVYNHLGNNDGMNLSAQTRRSKISKS 360

DB 301 LIGGDDFKSGQTRKMSVLVDPLVGAGIKPTYSVYNHLGNNDGMNLSAQTRRSKISKS 360

QY 361 NVVDDMV 367

DB 361 NVVDDMV 367

RESULT 2

T08436 inositol-3-phosphate synthase (EC 5.5.1.4) [similarity] - rape

C/Species: Brassica napus (rape)

C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004

C/Accession: T08436

R/Hussain, A.; Bourgeois, J.; Polvi, S.; Teang, E.; Keller, W.A.; George, F. submitted to the EMBL Data Library, August 1996

A/Reference number: Z16418

A/Accession: T08436

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-509 <HUS>

A/Cross-references: UNIPROT:Q96348; EMBL:U66307; NID:g1513227; PID:g1513228

C/function:

A/description: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo-

C/superfamily: myo-inositol-1-phosphate synthase

C/keywords: intramolecular lyase; isomerase

Query Match 13.1%; Score 67; DB 2; Length 509;

Best Local Similarity 100.0%; Pred. No. 1.7e-60;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTRFSKEISKS 360
|||||
Db 300 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTRFSKEISKS 359
|||||

QY 361 NVVDDMV 367
|||||
Db 360 NVVDDMV 366

RESULT 3

T04399

inositol-3-phosphate synthase (EC 5.5.1.4) - barley

C:Species: Hordeum vulgare (barley)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: T04399

R:Larson, S.R.; Raboy, V.

submitted to the EMBL Data Library, March 1998

A:Description: Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.

A:Reference number: Z14366

A:Accession: T04399

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-510 <LAR>

A:Cross-references: UNIPROT:065195; EMBL:AF056325; NID:g3152730; PIDN:AAC17133.1; PID:g3

A:Experimental source: cv. Harrington

C:Genetics:

A:Gene: INO1

A:Map position: 4

C:Function:

A:Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-inc

A:Note: first step

C:Superfamily: myo-inositol-1-phosphate synthase

C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 13.1%; Score 67; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.7e-60;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTRFSKEISKS 360
|||||
Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTRFSKEISKS 360
|||||

QY 361 NVVDDMV 367
|||||
Db 361 NVVDDMV 367

RESULT 4

T50021

inositol-3-phosphate synthase (EC 5.5.1.4) T31P16.160 [similarity] - Arabidopsis thaliana

N:Alternate names: protein T31P16.160

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004

C:Accession: T50021

R:Bevan, M.; Zimmermann, W.; Gruenisen, A.; Wambutt, R.; Kalicki, J.; Wohlmann, P.; Sm

submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25027

A:Accession: T50021

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-510 <BEV>

A:Cross-references: UNIPROT:Q9LX12; EMBL:AJ356332; GSPDB:GN00063; ATSP:T31P16.160

A:Experimental source: cultivar Columbia; BAC clone T31P16

C:Genetics:

A:Gene: ATSP:T31P16.160

A:Map position: 5

A:Introns: 63/2; 86/2; 131/3; 214/2; 290/1; 328/3; 387/3; 450/3; 471/3

C:Superfamily: myo-inositol-1-phosphate synthase

C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 13.1%; Score 67; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.7e-60;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTRFSKEISKS 360
|||||
Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTRFSKEISKS 360
|||||

QY 361 NVVDDMV 367
|||||
Db 361 NVVDDMV 367

RESULT 5

S60302

inositol-3-phosphate synthase (EC 5.5.1.4) - Spirodela polyrrhiza

C:Species: Spirodela polyrrhiza

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C:Accession: S60302

R:Smart, C.C.; Fleming, A.J.

Plant J. 4, 279-293, 1993

A:Title: A plant gene with homology to D-myo-inositol-3-phosphate synthase is rapidly an

A:Reference number: S60302; MUID:94035182; PMID:8220483

A:Accession: S60302

A:Molecule type: mRNA

A:Cross-references: UNIPROT:P42803; EMBL:Z11693; NID:g396209; PIDN:CAA77751.1; PID:g55864

C:Genetics:

A:Gene: cur1

C:Superfamily: myo-inositol-1-phosphate synthase

C:Keywords: intramolecular lyase; isomerase

Query Match 13.1%; Score 67; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.7e-60;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTRFSKEISKS 360
|||||
Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTRFSKEISKS 360
|||||

QY 361 NVVDDMV 367
|||||
Db 361 NVVDDMV 367

RESULT 6

D84610

probable myo-inositol 1-phosphate synthase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Aug-2004

C:Accession: D84610

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; H

W.; Koo, H.; Molnar, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.

enus, D.; Nierman, W.C.; White, J.A.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84610

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-510 <STO>

A:Cross-references: UNIPROT:Q38862; GB:AE002093; NID:g4567202; PIDN:AAD33618.1; GSPDB:GN

C:Genetics:

A:Gene: At2g22240

A:Map position: 2

C:Superfamily: Myo-inositol-1-phosphate synthase

Query Match 13.1%; Score 67; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.7e-60;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSTIVSYNHLGNNDGMNLSAPQTRFSKEISKS 360
|||||

Db 301 LIGGDDFKSGQTKMKSVLVDPLVAGAIKPTISVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
QY 361 NVVDDMV 367
Db 361 NVVDDMV 367

RESULT 7

T05017
Inositol-3-phosphate synthase (EC 5.5.1.4) T19P19.190 [similarity] - Arabidopsis thaliana
N/Alternate names: protein T19P19.190
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T05017
R/Sevan, M.; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Hohnsels, J.; Mewes, H.W.;
submitted to the Protein Sequence Database, April 1998
A/Reference number: Z15394
A/Accession: T05017
A/Molecule type: DNA
A/Residues: 1-511 <REV>
A/Cross-references: UNIPROT:P42801; EMBL:AL022605
A/Experimental source: cultivar Columbia; BAC clone T19P19
C/Genetics:
A/Map position: 4
A/Introns: 64/2; 87/2; 132/3; 215/2; 291/1; 329/3; 388/3; 451/3
A/Note: T19P19.190
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase; NAD

Query Match 13.1%; Score 67; DB 2; Length 511;
Best Local Similarity 100.0%; Pred. No. 1.7e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDPLVAGAIKPTISVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Db 302 LIGGDDFKSGQTKMKSVLVDPLVAGAIKPTISVSYNHLGNNDGMNLSAPQTFRSKEISKS 361
QY 361 NVVDDMV 367
Db 362 NVVDDMV 368

RESULT 8

T12438
Inositol-3-phosphate synthase (EC 5.5.1.4) - common ice plant
C/Species: Mesembryanthemum crystallinum (common ice plant)
C/Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C/Accession: T12438
R/Ishikawa, M.; Majumder, A.L.; Bornhauser, A.; Michalowski, C.B.; Jensen, R.G.; Bohmert
Plant J. 9, 537-548, 1996
A/Title: Coordinate transcriptional induction of myo-inositol metabolism during environm
A/Reference number: Z17518; M0ID:96208959; PMID:8624516
A/Accession: T12438
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: mRNA
A/Residues: 1-512 <ISH>
A/Cross-references: UNIPROT:Q40271; EMBL:U32511; NID:G975887; PIDN:AB03687.1; PID:G9758
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase

Query Match 13.1%; Score 67; DB 2; Length 512;
Best Local Similarity 100.0%; Pred. No. 1.7e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDPLVAGAIKPTISVSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Db 303 LIGGDDFKSGQTKMKSVLVDPLVAGAIKPTISVSYNHLGNNDGMNLSAPQTFRSKEISKS 362
QY 361 NVVDDMV 367
Db 363 NVVDDMV 369

RESULT 9

T01647
Inositol-3-phosphate synthase (EC 5.5.1.4) - maize
C/Species: Zea mays (maize)
C/Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
C/Accession: T01647
R/Larson, S.R.; Raboy, V.
submitted to the EMBL Data Library, March 1998
A/Description: linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.
A/Reference number: Z14366
A/Accession: T01647
A/Status: translated from GB/EMBL/DBDJ
A/Molecule type: mRNA
A/Residues: 1-510 <LAR>
A/Cross-references: UNIPROT:Q9FPK7; EMBL:AF056326; NID:G3108052; PIDN:AACT5756.1; PID:G31
A/Experimental source: strain Early ACR; leaf
C/Genetics:
A/Gene: INO1
C/Function:
A/Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-ino
A/Pathway: inositol biosynthesis
A/Note: NAD cofactor
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase; NAD

Query Match 9.8%; Score 50; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 5.9e-43;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMKSVLVDPLVAGAIKPTISVSYNHLGNNDGMNLSAPQ 350
Db 301 LIGGDDFKSGQTKMKSVLVDPLVAGAIKPTISVSYNHLGNNDGMNLSAPQ 350

RESULT 10

T10964
Inositol-3-phosphate synthase (EC 5.5.1.4) - kidney bean
N/Alternate names: 1L-myo-inositol 1-phosphate synthase
C/Species: Phaseolus vulgaris (kidney bean)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T10964
R/Wang, X.; Johnson, M.D.
submitted to the EMBL Data Library, October 1995
A/Reference number: Z17234
A/Accession: T10964
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: mRNA
A/Residues: 1-511 <WAN>
A/Cross-references: UNIPROT:Q41107; EMBL:U38920; NID:G1066282; PID:G1066283
A/Experimental source: strain Taylor's horticultural; root
C/Function:
A/Description: catalyzes reversible conversion of D-glucose 6-phosphate to 1L-myo-inositc
A/Pathway: myo-inositol biosynthesis
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase; NAD

Query Match 9.4%; Score 48; DB 2; Length 511;
Best Local Similarity 100.0%; Pred. No. 6.8e-41;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 DPLVAGAIKPTISVSYNHLGNNDGMNLSAPQTFRSKEISKS NVVDDMV 367
Db 321 DPLVAGAIKPTISVSYNHLGNNDGMNLSAPQTFRSKEISKS NVVDDMV 368

RESULT 11

T46317
Hypothetical protein DKFZp434A0612.1 - human
C/Species: Homo sapiens (man)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C/Accession: T46317
R/Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23035
A:Accession: T46317
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-430 <AAA>
A:Cross-references: UNIPROT:Q9NSU0; EMBL:AL137749
A:Experimental source: adult testis; clone DKP2p344A0612
C:Genetics:
A:Note: DKFZp344A0612.1
C:Superfamily: myo-inositol-1-phosphate synthase

Query Match 2.9%; Score 15; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 6.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 VIKKVPYVGSCKRA 398
DB 253 VIKKVPYVGSCKRA 267

RESULT 12

inositol-3-phosphate synthase (EC 5.5.1.4) - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18569; T20002
R:Ainscough, R.
submitted to the EMBL Data Library, December 1998
A:Reference number: Z18979
A:Accession: T18569
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-525 <WIL>
A:Cross-references: UNIPROT:Q18664; EMBL:AL033535; P1DN:CAA22132.1; CESP:VF13D12L.1
A:Experimental source: clone VF13D12L
R:Gajdasty, S.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z19209
A:Accession: T20002
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-525 <W12>
A:Cross-references: EMBL:269902; P1DN:CAA93771.1; GSPDB:GN00020; CESP:VF13D12L.1
A:Experimental source: clone C47D12
C:Genetics:
A:Gene: CESP:VF13D12L.1
A:Map position: 2
A:introns: 106/2; 287/1; 411/2
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase

Query Match 2.5%; Score 13; DB 2; Length 525;
Best Local Similarity 100.0%; Pred. No. 9.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 SIIVSYNHLGNNDG 343
DB 337 SIIVSYNHLGNNDG 349

RESULT 13

A30902
inositol-3-phosphate synthase (EC 5.5.1.4) [validated] - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein J0610; protein YJL153c
C:Species: *Saccharomyces cerevisiae*
C>Date: 18-Apr-1989 #sequence_revision 08-Sep-1995 #text_change 16-Aug-2004
C:Accession: S55160; B32209; S56935; S71644; A30827; A30902
R:Katsoulou, C.; Tzermita, M.; Alexandraki, D.
submitted to the EMBL Data Library, May 1995
A:Description: The complete sequence of a 40.7 kb segment located on the left arm of yeast hypothetical proteins.
A:Reference number: S55159
A:Accession: S55160

A:Molecule type: DNA
A:Residues: 1-555 <KAT>
A:Cross-references: EMBL:X87371; NID:G854542; PID:G854544
J:Dean-Johnson, M.; Henry, S.A.
U: Biol. Chem. 264, 1274-1283, 1989
A:Title: Biosynthesis of inositol in yeast. Primary structure of myo-inositol-1-phosphate
A:Reference number: A32209; MUID:89093118; PMID:2642902
A:Accession: B32209
A:Molecule type: DNA
A:Residues: 23-35, 'RL', 37-81, 'FE', 83-87, 'TRNYAHVWR', 88, 'QQW', 92-103, 'WPRYWRISTWWS', 116-1
V:SPQRLSFSFSAVL, <DEA>
A:Cross-references: EMBL:J04453
R:Katsoulou, C.; Tzermita, M.; Alexandraki, D.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56912
A:Accession: S56935
A:Molecule type: DNA
A:Residues: 1-555 <KAN>
A:Cross-references: EMBL:Z49428; NID:G1015570; PID:G1015571; MIPS:YJL153c
R:Katsoulou, C.; Tzermita, M.; Tavernarakis, N.; Alexandraki, D.
Yeast 12, 787-797, 1996
A:Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome X 1
of chromosome XI.
A:Reference number: S71643; MUID:96408771; PMID:8613765
A:Accession: S71644
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-555 <KAF>
A:Cross-references: EMBL:X87371; NID:G854542; P1DN:CAA60802.1; PID:G854544
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
C:Genetics:
A:Gene: SGD:INO1
A:Cross-references: SGD:S0003689; MIPS:YJL153c
A:Map position: 10L
C:Complex: homotetramer
C:Function: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo-
A:pathway: inositol biosynthesis
A:Note: requires NAD
C:Superfamily: Myo-inositol-1-phosphate synthase
C:Keywords: cytosol; homotetramer; intramolecular lyase; isomerase; NAD

Query Match 2.5%; Score 13; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 INGSPONTFVPGT 291
DB 315 INGSPONTFVPGT 327

RESULT 14

S45452
inositol-3-phosphate synthase (EC 5.5.1.4) - yeast (*Candida albicans*)
C:Species: *Candida albicans*
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S45452
R:Klig, L.S.; Zobel, P.A.; Devry, C.G.; Losberger, C.
Yeast 10, 789-800, 1994
A:Title: Comparison of INO1 gene sequences and products in *Candida albicans* and *Saccharom*
A:Reference number: S45452; MUID:95065381; PMID:7975896
A:Accession: S45452
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-520 <KLI>
A:Cross-references: UNIPROT:P42800; EMBL:L22737; NID:G413758; P1DN:AAA62849.1; PID:G69575;
A:Genetics:
A:Gene: INO1
A:Function: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo-
A:Note: requires NAD
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 2.4%; Score 12; DB 2; Length 520;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 INGSPTFTVPG 290
|||||
DB 285 INGSPTFTVPG 296

RESULT 15

B48357
Minor outer capsid protein - porcine rotavirus C (strain Cowden)
N:Alternate names: nonstructural protein NS26
C:Species: porcine rotavirus C
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: B48357
R:Bremont, M.; Chabanne-Vautherot, D.; Cohen, J.
Arch. Virol. 130, 85-92, 1993
A:Title: Sequence analysis of three non structural proteins of a porcine group C (Cowden)
A:Reference number: A48357; MUID:93277387; PMID:8389118
A:Accession: B48357
A:Molecule type: genomic RNA
A:Residues: 1-210 <BRS>
A:Cross-references: UNIPROT:P36358
A:Note: sequence extracted from NCBI backbone (NCBIN:132693, NCBI:P132694)
C:Genetics:
A:Map position: segment 10
C:Superfamily: rotavirus minor outer capsid protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:30,120/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.6%; Score 8; DB 1; Length 210;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 313 KMKSVLVD 320
|||||
DB 198 KMKSVLVD 205

Search completed: June 7, 2005, 17:10:36
Job time : 28.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:49:04 ; Search time 86 Seconds
(without alignments)
3036.749 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 510

Sequence: 1 MEIENFKVCECPNVKXTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : UniProt 03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	510	2	094C02
2	376	73.7	510	2	094K03
3	92	18.0	472	2	094G22
4	75	14.7	510	2	07XJC0
5	70	13.7	510	1	IN01_TOBAC
6	69	13.5	443	2	08H1B7
7	69	13.5	510	1	IN01_NICPA
8	67	13.1	364	2	09AMG8
9	67	13.1	507	1	IN01_CITPA
10	67	13.1	509	2	09AR12
11	67	13.1	510	1	IN01_BRANA
12	67	13.1	510	1	IN01_HORVU
13	67	13.1	510	1	IN01_ORYSA
14	67	13.1	510	1	IN01_SERSTIN
15	67	13.1	510	1	IN01_SEIPO
16	67	13.1	510	1	IN01_WHEAT
17	67	13.1	510	1	IN02_ARATH
18	67	13.1	510	1	IN03_ARATH
19	67	13.1	510	2	0944C3
20	67	13.1	510	2	09AV99
21	67	13.1	511	1	IN01_ARATH
22	67	13.1	512	1	IN01_MESCR
23	50	9.8	510	1	IN01_MAIZE
24	48	9.4	511	1	IN01_PHAUV
25	47	9.2	164	2	09FT40
26	43	8.4	512	2	07XZE6
27	38	7.5	186	2	042548
28	36	7.1	409	2	0845A1
29	31	6.1	509	2	0885N2
30	29	5.7	565	1	IN01_DROME
31	27	5.3	135	2	09AT03

32	27	5.3	166	2	09FUP2	09FUP2 lycopersico
33	23	4.5	150	2	095U04	095U04 brachioleto
34	23	4.5	560	2	06DDT1	06DDT1 xenopus lae
35	23	4.5	563	2	07ZXU0	07ZXU0 xenopus lae
36	21	4.1	561	2	07P2B9	07P2B9 anopheles g
37	17	3.3	526	2	07Z525	07Z525 homo sapien
38	17	3.3	557	2	09JHU9	09JHU9 m myo-1nos1
39	17	3.3	558	2	06NXT5	06NXT5 homo sapien
40	17	3.3	558	2	09H2Y2	09H2Y2 homo sapien
41	17	3.3	558	2	09NPH2	09NPH2 homo sapien
42	17	3.3	558	2	09NHW7	09NHW7 homo sapien
43	16	3.1	220	2	06UP00	06UP00 aspergillus
44	16	3.1	537	2	087IU5	087IU5 neurospora
45	15	2.9	181	2	09BT65	09BT65 homo sapien

ALIGNMENTS

RESULT 1	ID	Q94C02	PRELIMINARY;	PRT;	510 AA.
AC	Q94C02;				
DT	01-DEC-2001 (TrEMBLrel. 19, Created)				
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)				
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)				
DE	Myo-inositol-1-phosphate synthase (EC 5.5.1.4).				
OS	Glycine max (soybean).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.				
OX	NCBI_TaxID=3847;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Hitz W.D., Carlson T.J., Kerr P., Sebastian S.;				
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.				
RT	EMBL: AY038802; AK72098.1; -				
DR	HSSP; P12986; 1P1H.				
DR	GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.				
DR	GO; GO:0016853; F:isomerase activity; IEA.				
DR	GO; GO:0006021; P:myo-inositol biosynthesis; IEA.				
DR	GO; GO:0008654; P:phospholipid biosynthesis; IEA.				
DR	InterPro: IPR002587; Inos-1-P synth.				
DR	Pfam: PF01658; Inos-1-P synth; 1.				
KW	Isomerase.				
SQ	SEQUENCE 510 AA; 56475 MW; A1B4C7F1643918E CRC64;				
Query Match	100.0%; Score 510; DB 2; Length 510;				
Best Local Similarity	100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;				
Matches 510; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	MEIENFKVCECPNVKXTETETI	QSVVYVETTEL	VHNRNGTYQVI	IVPKSVKYEFTKNIHVP 60
DB	1	MEIENFKVCECPNVKXTETETI	QSVVYVETTEL	VHNRNGTYQVI	IVPKSVKYEFTKNIHVP 60
QY	61	KLGWMLVGGNGSGTLGGVIANR	BEGISWATKDKI	QOANYGSGSL	TOASAIIVGSGQEE 120
DB	61	KLGWMLVGGNGSGTLGGVIANR	BEGISWATKDKI	QOANYGSGSL	TOASAIIVGSGQEE 120
QY	121	IYAPFKSLIPVNPDPDIFVGG	MDISNMNLAD	MAKAVPDID	LOQLRPMESMLPLPGI 180
DB	121	IYAPFKSLIPVNPDPDIFVGG	MDISNMNLAD	MAKAVPDID	LOQLRPMESMLPLPGI 180
QY	181	YDPDFIAANOEERANVIGKT	QEOVOQI	IKDKIKAFK	EATKVDKVVVLTANTERYSNLV 240
DB	181	YDPDFIAANOEERANVIGKT	QEOVOQI	IKDKIKAFK	EATKVDKVVVLTANTERYSNLV 240
QY	241	VLANTNTENLLAANDRNEA	ISPSLTVAI	ACMENVPI	NGSPQTFPVGILDLAIART 300
DB	241	VLANTNTENLLAANDRNEA	ISPSLTVAI	ACMENVPI	NGSPQTFPVGILDLAIART 300
QY	301	LIGGDPFSGQTKMSVVD	FLVAGIK	IPSTISV	VNHLGNDGNMLSAPOFRSKEISK 360
DB	301	LIGGDPFSGQTKMSVVD	FLVAGIK	IPSTISV	VNHLGNDGNMLSAPOFRSKEISK 360

Db	301	LIGGDDFFSGGQTKMKMSVLVDPLVGAGIKRTSIVSYHNLGNDGMNLSAQOTPSRSKISKS	360
Qy	361	NVVDMDVNSNALIYPERGEHPDHVVYIKYVPYVGDSSGRAMDEYTSSEIFMGGSSTIVLHNTC	420
Db	361	NVVDMDVNSNALIYPERGEHPDHVVYIKYVPYVGDSSGRAMDEYTSSEIFMGGSSTIVLHNTC	420
Qy	421	EDSLLAAPILDLVLLAEISTRIEFAENBEGKPHSHPAVATIIISYITKAPLVPRGTPVYN	480
Db	421	EDSLLAAPILDLVLLAEISTRIEFAENBEGKPHSHPAVATIIISYITKAPLVPRGTPVYN	480
Qy	481	ALSKORAMLENTMRACVGLAPENNMTLEYSK	510
Db	481	ALSKORAMLENTMRACVGLAPENNMTLEYSK	510

RESULT 2		
ID	Q94KU3	PRELIMINARY; PRT; 510 AA.
AC	Q94KU3;	
DT	01-DEC-2001 (TREMBLrel. 19, Created)	
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)	
DE	Myo-inositol-3-phosphate synthase.	
GN	Name=MIPS;	
OS	Glycine max (soybean).	
OC	Euparieta; Vitidiplophidae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;	
OC	eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseolaceae; Glycine.	
OX	NCBI_TaxID=3847;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=21196082; PubMed=11299373; DOI=10.1104/PP.125.4.1941;	
RA	Hegeman C.E., Good L.L., Grabau E.A.;	
RT	"Expression of D-myo-inositol-3-phosphate synthase in soybean.	
RL	Plant Physiol. 125:1941-1948 (2001).	
RL	EMBL; AF293970; AAK49896.1; -.	
DR	HGSP; P11986; 1PH.	
DR	GO; GO:0004513; Pinositol-3-phosphate synthase activity; IEA.	
DR	GO; GO:0006021; P:myo-inositol biosynthesis; IEA.	
DR	GO; GO:0008654; P:phospholipid biosynthesis; IEA.	
DR	InterPro; IPR002587; Inos-1_P_synth.	
DR	Pfam; PF01658; Inos-1_P_synth; 1.	
SO	SEQUENCE 510 AA; 56506 MW; DE4F3DD7DC6F370 CRC64;	

Query Match	73.7%;	Score 376;	DB 2;	Length 510;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 476;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

QY	1	MEIENFKCEPNVKYETETISQSYANNETTELHENNGTYQWLVKPSVYEEKRTNHP	60
Db	1	MEIENFKCEPNVKYETETISQSYANNETTELHENNGTYQWLVKPSVYEEKRTNHP	60
QY	61	KLGWMLVGWGNNGSTLTGGLVIANREGISWATQDKIQOANYFGSLTQASAIRGSGOGE	120
Db	61	KLGWMLVGWGNNGSTLTGGLVIANREGISWATQDKIQOANYFGSLTQASAIRGSGOGE	120
QY	121	IYAPFKSLPLPMVNPDDIVFGGMDISNMNLLADAMARAKVPDIDLOKOLRPMESMLPLPGI	180
Db	121	IYAPFKSLPLPMVNPDDIVFGGMDISNMNLLADAMARAKVPDIDLOKQWRPMESSLPLPGI	180
QY	181	YDPDFIAANOEBRANNVIKGTQEOVOQIIKDIKAFKEATKVDKVVVLMTANTERSYNLV	240
Db	181	YDPDFIAANOEBRANNVIKGTQEOVOQIIKDIKAFKEATKVDKVVVLMTANTERSYNLV	240
QY	241	VELNDTMEMLLAAVDNNEAEISSTLYAIACWENNPFINGSPONTFVPGILDLAIARNT	300
Db	241	VELNDTMEMLLAAVDNNEAEISSTLYAIACWENNPFINGSPONTFVPGILDLAIARNT	300
QY	301	LIGGDDPKSGQRTKMSVLVDLVLGAGIKPRTISYSYNHLGANNDDGNLSAPOTPRSEIKSKS	360
Db	301	LIGGDDPKSGQRTKMSVLVDLVLGAGIKPRTISYSYNHLGANNDDGNLSAPOTPRSEIKSKS	360

Qy	361	NVYDDMVNSNALIYEEGGEHPDHVVIKVPYVYGS KRAMDEYTSGEIIMGSGSTIVLHNTC	420
Db	361	NVYDDMVNSNALIYEEGGEHPDHVVIKVPYVYGS KRAMDEYTSGEIIMGSGSTIVLHNTC	420
Qy	421	EDSLAAPITLIDLVYLAELSTRIEKKAENECKFHSFHFVAITLSYVLKAPLVPKPGTP	477
Db	421	EDSLAAPITLIDLVYLAELSTRIEKKAENECKFHSFHFVAITLSYVLKAPLVPKPGTP	477

```

RESULT 3
094G22
ID 094G22 PRELIMINARY; PRT; 472 AA.
AC 094G22;
DT 01-DEC-2001 (TREMBLrel_19, Created)
DT 01-DEC-2001 (TREMBLrel_19, Last sequence update)
DT 01-JUN-2003 (TREMBLrel_24, Last annotation update)
DE 1L-myo-inositol-1-phosphate synthase (Fragment) .
OS Phaseolus vulgaris (Kidney bean) (French bean) .
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids
OC eusteroids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus
NM NCBI_TaxID=3885;
NM [1]
RP SEQUENCE FROM N.A.
RA Johnson M.D., Lackey K.H., Pope P.M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF282263; AAK69514.1; -.
DR HSSP; P11986; 191H.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P_synth.
FT Pfam; PF01658; Inos-1-P_synth; 1.
SQ NON_TER 472 472
SC SEQUENCE 472 AA; 237204E1A370560F CRC64;

```

Query Match 18.0%; Score 92; DB 2; Length 472;
Best Local Similarity 100.0%; Pred. No. 7e-87;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	301	LIIGDDPKSQTOMKSVLVDPLVGAGIKRTSVSVNHLGNDGNMLSPQTFRSEIKS	360
Db	301	LIIGDDPKSQTOMKSVLVDPLVGAGIKRTSVSVNHLGNDGNMLSPQTFRSEIKS	360
QY	361	NVYDDMYSNAIIYEPGEHDPHVVIKYPYV	392
Db	361	NVYDDMYSNAIIYEPGEHDPHVVIKYPYV	392

```

RESULT 4
Q7XJCO
ID Q7XJCO PRELIMINARY; PRT; 510 AA.
AC Q7XJCO;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Myo-inositol-1-phosphate synthase INOL.
OS Xerophyta viscosa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
OC Spermatophytes; Magnoliophyta; Liliopsida; Velloziaceae; Xerophyta
OX NCBI_TaxID=96708;
RN [1]
RP SEQUENCE FROM N.A.
RA Majee M., Majumder A.N.L., Mundree S.G.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY323824; AAP85531.1; -.
DR HSSP; P11986; IJXI.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:inositol lipid biosynthesis; IEA.
DR InterPro; IPR02587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
SQ SEQUENCE 510 AA; 56534 MW; 5F92212851115A2A CRC64;

```

Query Match 14.7%; Score 75; DB 2; Length 510;
 Best Local Similarity 100.0%; Pred. No. 4,7e-69;
 Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 SNAIYEPGEHPDHVVVVKYPPVYVDSKRAMDEYTSSEIFMGKSTIVLHNTCEDSLAAPP 428
 DB 369 SNAIYEPGEHPDHVVVVKYPPVYVDSKRAMDEYTSSEIFMGKSTIVLHNTCEDSLAAPP 428
 QY 429 IILDVVLAEIESTRI 443
 DB 429 IILDVVLAEIESTRI 443

RESULT 5

INOL TOBAC STANDARD; PRT; 510 AA.
 AC 09LW96;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (MI-1-P synthase) (IPS).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; asterids; Solanales; Solanaceae; Nicotiana.
 NCBI_TaxID=4097;
 RX MEDLINE=20399434; PubMed=10945337;
 RA Hara K., Yagi M., Koizumi N., Kusano T., Sano H.;
 RT "Screening of wound-responsive genes identifies an immediate-early expressed gene encoding a highly charged protein in mechanically wounded tobacco plants."
 RL Plant Cell Physiol. 41:684-691(2000).
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC EMBL: AB009881; BAA95788.1; -.
 DR HSSP: P11986; 1PKF.
 DR InterPro: IPR002587; Inos-1_P_synth.
 DR Pfam: PF01658; Inos-1_P_synth; 1.
 DR Inositol biosynthesis; Isomerase; NAD; phospholipid biosynthesis.
 SQ SEQUENCE 510 AA; 56369 MW; 4BA6FDDA5DF6D4D CRC64;

Query Match 13.7%; Score 70; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 8.1e-64;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 RNTLIGDDPKSGQTKMSVLVDFLVGAGIKPTSIYSYVNLGNNDGMSAPQTFPSKEI 357
 DB 298 RNTLIGDDPKSGQTKMSVLVDFLVGAGIKPTSIYSYVNLGNNDGMSAPQTFPSKEI 357
 QY 358 SKSNVVDNV 367
 DB 358 SKSNVVDNV 367

RESULT 6

08H1B7 PRELIMINARY; PRT; 443 AA.
 AC 08H1B7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Myo-inositol phosphate synthase (fragment).
 OS Lolium perenne (Perennial ryegrass).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Poaceae; Lolium.
 NCBI_TaxID=4522;
 RX SEQUENCE FROM N.A.
 RA Amlard V., Prud'homme M.-P., Le Dantec C.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY154382; AAN52772.1; -.
 DR HSSP: P11986; 1PKF.
 DR GO: GO:0004512; P:inositol-3-phosphate synthase activity; IEA.
 DR GO: GO:0006021; P:myo-inositol biosynthesis; IEA.
 DR GO: GO:0006654; P:phospholipid biosynthesis; IEA.
 DR InterPro: IPR002587; Inos-1-P_synth.
 DR Pfam: PF01658; Inos-1-P_synth; 1.
 FT NON_TER 443

Query Match 13.5%; Score 69; DB 2; Length 443;
 Best Local Similarity 100.0%; Pred. No. 8e-63;
 Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 375 EGEHPDHVVVVKYPPVYVDSKRAMDEYTSSEIFMGKSTIVLHNTCEDSLAAPP 434
 DB 375 EGEHPDHVVVVKYPPVYVDSKRAMDEYTSSEIFMGKSTIVLHNTCEDSLAAPP 434
 QY 435 LLAELSTRI 443
 DB 435 LLAELSTRI 443

RESULT 7

INOL NICPA STANDARD; PRT; 510 AA.
 AC 09SST4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (MI-1-P synthase) (IPS).
 GN Name=INPS1;
 OS Nicotiana paniculata.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; asterids; Solanales; Solanaceae; Nicotiana.
 NCBI_TaxID=62141;
 RX SEQUENCE FROM N.A.
 RA Hashimoto A., Yamada S., Komori T.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC EMBL: AB009881; BAA95788.1; -.
 DR HSSP: P11986; 1PKF.
 DR InterPro: IPR002587; Inos-1_P_synth.
 DR Pfam: PF01658; Inos-1_P_synth; 1.
 DR Inositol biosynthesis; Isomerase; NAD; phospholipid biosynthesis.
 SQ SEQUENCE 510 AA; 56369 MW; 4BA6FDDA5DF6D4D CRC64;

```
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL: AB032073; BAA84084.1; -.
DR HSSP: P11986; 1PIH.
DR InterPro: IPR002587; Inos-1-P_synth.
DR Pfam: PF01658; Inos-1-P_synth; 1.
DR Inositol biosynthesis; Isomerase; NAD; phospholipid biosynthesis.
KW SEQUENCE 510 AA; 56385 MW; 415B81C27A267666 CRC64;
SQ
Query Match 13.1%; Score 69; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 9.1e-63;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 299 NTLLGGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNNHGNNDGMNLSAPQTFRSKEIS 358
DB 299 NTLLGGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNNHGNNDGMNLSAPQTFRSKEIS 358
QY 359 KSNVVDMMV 367
DB 359 KSNVVDMMV 367
RESULT 8
Q9AWG8 PRELIMINARY; PRT; 364 AA.
AC Q9AWG8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Myo-inositol-1-phosphate synthase.
OS Actinidia arguta.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Ericales; Actinidaceae; Actinidia.
OX NCBI_TaxID=64478;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Young mature leaf;
RA Klages K., Fitzgerald A., Moodie M.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY005128; ANP97409.1; -.
DR HSSP: P11986; 1PIK.
DR GO: GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO: GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO: GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro: IPR002587; Inos-1-P_synth.
DR Pfam: PF01658; Inos-1-P_synth; 1.
SQ SEQUENCE 364 AA; 40246 MW; 1CB64D6FF78127D CRC64;
Query Match 13.1%; Score 67; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 8.4e-61;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 301 LIGGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNNHGNNDGMNLSAPQTFRSKEIS 360
DB 155 LIGGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNNHGNNDGMNLSAPQTFRSKEIS 214
QY 361 NVVDDMMV 367
DB 215 NVVDDMMV 221
RESULT 9
ID INOI_CITPA STANDARD; PRT; 507 AA.
AC P42802;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
synthase) (Mt-1-P synthase) (IPS).
OS Citrus paradisi (Grapefruit).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=37656;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=leaf;
RA MEDLINE=95148748; PubMed=7846170; DOI=10.1104/pp.106.4.1689;
RX Abu-Abied M., Holland D.;
RT "The gene c-inol from Citrus paradisi is highly homologous to tur1 and
inol from yeast and Spirodela encoding for myo-inositol phosphate
synthase."
RT Plant Physiol. 106:1689-1689 (1994).
RL -I- CATALYTIC ACTIVITY: D-Glucose 6-phosphate = 1D-myo-inositol 3-
phosphate.
CC -I- COPOLYMER: NAD.
CC -I- PATHWAY: Inositol biosynthesis.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (potential).
CC -I- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL: Z22632; CAA83565.1; -.
DR PIR: S52648; S52648.
DR HSSP: P11986; 1PIK.
DR InterPro: IPR002587; Inos-1-P_synth.
DR Pfam: PF01658; Inos-1-P_synth; 1.
KW Inositol biosynthesis; Isomerase; NAD; phospholipid biosynthesis.
SQ SEQUENCE 507 AA; 56334 MW; 45D78928991BEDF8 CRC64;
Query Match 13.1%; Score 67; DB 1; Length 507;
Best Local Similarity 100.0%; Pred. No. 1.1e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 301 LIGGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNNHGNNDGMNLSAPQTFRSKEIS 360
DB 301 LIGGDDPKSGQTKMKSVLVDFLVGAGIKPTSIYSYNNHGNNDGMNLSAPQTFRSKEIS 360
QY 361 NVVDDMMV 367
DB 361 NVVDDMMV 367
RESULT 10
Q9ARI2 PRELIMINARY; PRT; 509 AA.
AC Q9ARI2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Myo-inositol 1-phosphate synthase.
OS Avicennia marina (Grey mangrove).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Acanthaceae; Acanthaceae incertae sedis; Avicennia.
OX NCBI_TaxID=82927;
RN [1]
RP SEQUENCE FROM N.A.
RA Jitresh M.N., Parani M., Parida A.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY028259; AAK21969.1; -.
DR HSSP: P11986; 1PIK.
DR GO: GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO: GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO: GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro: IPR002587; Inos-1-P_synth.
DR Pfam: PF01658; Inos-1-P_synth; 1.
```

SQ SEQUENCE 509 AA; 55978 MW; 23C8D354BAF3BD0F CRC64;
 Query Match 13.1%; Score 67; DB 2; Length 509;
 Best Local Similarity 100.0%; Pred. No. 1.1e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 DB 300 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSEIKS 359
 QY 361 NVVDDMV 367
 DB 360 NVVDDMV 366

RESULT 11
 INOI BRANA STANDARD; PRT; 510 AA.
 ID INOI BRANA
 AC 096348;
 DT 16-OCT-2001 (Rel. 40, Last Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (MI-1-P synthase) (IPS).
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Euroside II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hussein A., Bourgeois J., Polvi S., Teang E., Keller W.A., Georges F.;
 RT "Cloning of a full length cDNA encoding myo-inositol 1-phosphate synthase from Brassica napus."
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U66307; AAB06756.2; -.
 DR PIR: T08436; T08436.
 DR HSSP: P11986; 1PIU.
 DR InterPro: IPR002587; Inos-1-P synth.
 DR Pfam: PF01658; Inos-1-P synth. 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 SQ SEQUENCE 510 AA; 56377 MW; A40EB6558D80739 CRC64;

Query Match 13.1%; Score 67; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1.1e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 DB 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367

RESULT 12

INOI HORVU STANDARD; PRT; 510 AA.
 ID INOI HORVU
 AC 065195;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (MI-1-P synthase) (IPS).
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. Harrington;
 RA Larson S.R., Raboy V.;
 RT "Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes."
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF056325; AAC17133.1; -.
 DR PIR: T04399; T04399.
 DR HSSP: P11986; 1PIU.
 DR InterPro: IPR002587; Inos-1-P synth.
 DR Pfam: PF01658; Inos-1-P synth. 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 SQ SEQUENCE 510 AA; 56173 MW; EA63138121692724 CRC64;

Query Match 13.1%; Score 67; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1.1e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 DB 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSTIVSYNHLGNNDGMNLSAPQTRSKSEIKS 360
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367

RESULT 13
 INOI ORYSA STANDARD; PRT; 510 AA.
 ID INOI ORYSA
 AC 064437;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (MI-1-P synthase) (IPS).
 GN Name=INOI;
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriocaridaceae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=cnv. Japonica / Kamenoo;
 RA Yoshida K.T., Wada T., Koyama H., Mizobuchi-Fukunaka R., Naito S.;
 RT "Expression of myo-inositol 1-phosphate synthase gene and phytin
 accumulation during seed development in rice";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
 phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB012107; BAA25729.1; -.
 DR HSSP; P11986; 1PIH.
 DR Gramene; 064437; -.
 DR InterPro; IPR002587; Inos-1-P synth.
 DR Pfam; PF01658; Inos-1-P synth. 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 KW SEQUENCE 510 AA; 56216 MW; 2580220DD871A80 CRC64;

Query Match 13.1%; Score 67; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1.1e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGDDPFGSGQTKMKSVLVDFVGAGIKPTISIVSYNHLGNNDGNMLSAPOFRSKEISKS 360
 DB 301 LIGDDPFGSGQTKMKSVLVDFVGAGIKPTISIVSYNHLGNNDGNMLSAPOFRSKEISKS 360
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367

RESULT 14
 INOI SESIN STANDARD; PRT; 510 AA.
 AC Q9FYV1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
 synthase) (MT-1-P synthase) (IPS).
 OS Sesamum indicum (Oriental sesame) (Gingelly).
 CC Burkaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 CC Lamiales; Lamiales; Pedaliaceae; Sesamum.
 CC NCB1_Taxid=4182;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Seed;
 RA Jin U.-H., Chung C.-H.;
 RT "Characterization and functional analysis of a myo-inositol 1-
 phosphate synthase cDNA from sesame (Sesamum indicum L.) seeds";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
 phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF284065; AAC01148.1; -.
 DR HSSP; P11986; 1PIH.
 DR InterPro; IPR002587; Inos-1-P synth.
 DR Pfam; PF01658; Inos-1-P synth. 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 KW SEQUENCE 510 AA; 56234 MW; 88D75376CE873401F CRC64;

Query Match 13.1%; Score 67; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1.1e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGDDPFGSGQTKMKSVLVDFVGAGIKPTISIVSYNHLGNNDGNMLSAPOFRSKEISKS 360
 DB 301 LIGDDPFGSGQTKMKSVLVDFVGAGIKPTISIVSYNHLGNNDGNMLSAPOFRSKEISKS 360
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367

RESULT 15
 INOI SPIPO STANDARD; PRT; 510 AA.
 AC P42803;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
 synthase) (MT-1-P synthase) (IPS).
 GN Name=TUR1;
 OS Spirodela polyrrhiza (Giant duckweed).
 CC Burkaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Araceae; Lemnoidae;
 CC Spirodela.
 CC NCB1_Taxid=29656;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94035182; PubMed=8220483;
 RA Smart C.C., Fleming A.U.;
 RT "A plant gene with homology to D-myo-inositol-3-phosphate synthase is
 RT rapidly and spatially up-regulated during an abscisic-acid-induced
 RT morphogenic response in Spirodela polyrrhiza";
 RL Plant J. 4:279-293(1993).
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
 phosphate.
 CC -1- COFACTOR: NAD.
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- INDUCTION: By abscisic acid (ABA).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z11693; CAA7751.1; -.
 DR PIR; S60302; S60302.
 DR HSSP; P11986; 1PIH.
 DR InterPro; IPR002587; Inos-1-P synth.
 DR Pfam; PF01658; Inos-1-P synth. 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.

SQ SEQUENCE 510 AA; 56385 MW; 2D56D366FC5E03C CRC64;

Query Match 13.1%; Score 67; DB 1; Length 510;

Best Local Similarity 100.0%; Pred.No. 1.1e-60;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDPKSGQTMKSVLVDFLVGAGIKPTSIYSYNNHLGNNDSMTLSAPOTFRSKEISKS 360

DB 301 LIGGDDPKSGQTMKSVLVDFLVGAGIKPTSIYSYNNHLGNNDSMTLSAPOTFRSKEISKS 360

QY 361 NVVDDMV 367

DB 361 NVVDDMV 367

Search completed: June 7, 2005, 17:09:33

Job time : 88 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 17:00:20 ; Search time 31 Seconds
(without alignments)
1228.098 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 510
Sequence: 1 MEIENPKVCECPVKYETETI.....NIMRACVGLAPENNMILEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
 - 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
 - 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
 - 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
 - 5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
 - 6: /cgn2_6/prodata/1/iaa/backfile01.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	13.3	510	4	US-09-727-628-2
2	67	13.1	510	3	US-09-118-442-11
3	67	13.1	510	3	US-09-677-064-11
4	13	2.5	533	4	US-09-734-237B-73
5	13	2.5	534	4	US-09-734-237B-75
6	12	2.4	525	4	US-09-248-796A-17234
7	9	1.8	1481	4	US-09-231-899-70
8	7	1.4	55	4	US-09-640-211A-2116
9	7	1.4	79	4	US-09-107-433-4073
10	7	1.4	93	4	US-09-502-540-15484
11	7	1.4	134	4	US-09-270-767-34609
12	7	1.4	134	4	US-09-270-767-49826
13	7	1.4	160	4	US-09-957-641A-19
14	7	1.4	191	4	US-09-252-991A-24024
15	7	1.4	224	2	US-08-372-255-16
16	7	1.4	224	5	PCT-US95-08565-16
17	7	1.4	254	4	US-09-252-991A-21339
18	7	1.4	269	4	US-09-902-540-16003
19	7	1.4	273	3	US-08-235-836C-142
20	7	1.4	273	3	US-08-235-836C-144
21	7	1.4	284	4	US-09-914-259-62
22	7	1.4	284	4	US-10-164-595-32
23	7	1.4	295	4	US-09-270-767-57050
24	7	1.4	295	4	US-09-270-767-57401
25	7	1.4	306	4	US-09-107-532A-5995
26	7	1.4	308	4	US-09-489-039A-11598
27	7	1.4	322	4	US-09-565-501A-112

28	7	1.4	322	4	US-09-639-206A-112	Sequence 112, App
29	7	1.4	322	4	US-09-874-923-112	Sequence 112, App
30	7	1.4	338	4	US-09-583-110-4209	Sequence 4209, Ap
31	7	1.4	340	4	US-09-489-039A-9804	Sequence 9804, Ap
32	7	1.4	360	4	US-09-489-039A-9936	Sequence 9936, Ap
33	7	1.4	361	4	US-09-710-279-160	Sequence 160, App
34	7	1.4	364	4	US-09-902-540-16435	Sequence 16435, A
35	7	1.4	381	4	US-09-254-776B-53	Sequence 53, Appl
36	7	1.4	388	4	US-09-949-016-7631	Sequence 7631, Ap
37	7	1.4	399	4	US-09-543-681A-5325	Sequence 5325, Ap
38	7	1.4	443	4	US-09-270-767-46616	Sequence 46616, A
39	7	1.4	443	4	US-09-949-016-11613	Sequence 11613, A
40	7	1.4	446	4	US-09-252-991A-22844	Sequence 22844, A
41	7	1.4	456	4	US-09-328-352-6946	Sequence 6946, Ap
42	7	1.4	467	4	US-09-489-039A-12726	Sequence 12726, A
43	7	1.4	472	4	US-09-489-039A-8121	Sequence 8121, Ap
44	7	1.4	472	4	US-09-134-000C-3893	Sequence 3893, Ap
45	7	1.4	548	4	US-09-902-540-11870	Sequence 11870, A

ALIGNMENTS

```

RESULT 1
US-09-727-628-2
; Sequence 2, Application US/09727628
; Patent No. 6791013
; GENERAL INFORMATION:
; APPLICANT: Armstrong, Katherine
; APPLICANT: Hey, Timothy D
; APPLICANT: Folkerts, Otto
; APPLICANT: Smith, Kelley A
; APPLICANT: Hopkins, Nicole L
; TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
; FILE REFERENCE: 50597
; CURRENT APPLICATION NUMBER: US/09/727,628
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/168,612
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-727-628-2
Query Match      13.3%; Score 68; DB 4; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.1e-62;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      376 PGEHPDVVVVKKVPYVSGSKRAMDEYTSIFMGKSTVLNHTCDSLLAAPTIDLVL 435
      |||||
Db      376 PGEHPDVVVVKKVPYVSGSKRAMDEYTSIFMGKSTVLNHTCDSLLAAPTIDLVL 435
      |||||
OY      436 LABELSTRI 443
      |||||
Db      436 LABELSTRI 443
      |||||
RESULT 2
US-09-118-442-11
; Sequence 11, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catv, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706

```

```

; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 510
; TYPE: PRF
; ORGANISM: Zea mays
US-09-118-442-11
```

```

Query Match      13.1%; Score 67; DB 3; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.4e-61;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 301 LGGDDFKSGQTKMSVLVDPLVGAGIKPTSIYSYHNLGNNDGMNLSAPQTFRSKXISKS 360
Db 301 LGGDDFKSGQTKMSVLVDPLVGAGIKPTSIYSYHNLGNNDGMNLSAPQTFRSKXISKS 360
```

```
QY 361 NVVDDMV 367
Db 361 NVVDDMV 367
```

```

RESULT 3
US-09-677-064-11
; Sequence 11, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phyate Metabolism in
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 510
; TYPE: PRF
; ORGANISM: Zea mays
US-09-677-064-11
```

```

Query Match      13.1%; Score 67; DB 3; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.4e-61;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 301 LGGDDFKSGQTKMSVLVDPLVGAGIKPTSIYSYHNLGNNDGMNLSAPQTFRSKXISKS 360
Db 301 LGGDDFKSGQTKMSVLVDPLVGAGIKPTSIYSYHNLGNNDGMNLSAPQTFRSKXISKS 360
```

```
QY 361 NVVDDMV 367
Db 361 NVVDDMV 367
```

```

RESULT 4
US-09-734-237B-73
```

```

; Sequence 73, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 533
; TYPE: PRF
; ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-73
```

```

Query Match      2.5%; Score 13; DB 4; Length 533;
Best Local Similarity 100.0%; Pred. No. 9.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 279 INGSPONTFVPG 291
Db 293 INGSPONTFVPG 305
```

```

RESULT 5
US-09-734-237B-75
; Sequence 75, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 534
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic protein derived from Saccharomyces cerevisiae myo-inositol-1-phosphate synthase, having a glycine residue inserted after
; OTHER INFORMATION: the initiating methionine
US-09-734-237B-75
```

```

Query Match      2.5%; Score 13; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 9.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 279 INGSPONTFVPG 291
Db 294 INGSPONTFVPG 306
```

```

RESULT 6
US-09-248-796A-17234
; Sequence 17234, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
```

```

; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17234
; LENGTH: 525
; TYPE: PR1
; ORGANISM: Candida albicans
US-09-248-796A-17234

Query Match
Best Local Similarity 100.0%; Score 12; DB 4; Length 525;
Pred. No. 0.0011; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 279 INGSPTFTVPG 290
Db 290 INGSPTFTVPG 301

RESULT 7
US-09-231-899-70
; Sequence 70, Application US/09231899
; Patent No. 656583
; GENERAL INFORMATION:
; APPLICANT: Lassner, Michael
; APPLICANT: Metz, James G
; APPLICANT: Faciollet, Daniel
; TITLE OF INVENTION: SCHIZOCHYTRIUM PKS GENES
; FILE REFERENCE: CGNE.131.02US
; CURRENT APPLICATION NUMBER: US/09/231, 899
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/048,650
; EARLIER FILING DATE: 1997-06-04
; EARLIER APPLICATION NUMBER: 09/090,793
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 70
; LENGTH: 1481
; TYPE: PR1
; ORGANISM: Schizochytrium aggregatum
US-09-231-899-70

Query Match
Best Local Similarity 100.0%; Score 9; DB 4; Length 1481;
Pred. No. 4; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 248 ENLLAAVDR 256
Db 1092 ENLLAAVDR 1100

RESULT 8
US-09-640-211A-2116
; Sequence 2116, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2116
; LENGTH: 55
```

```

; TYPE: PR1
; ORGANISM: Eucalyptus grandis
US-09-640-211A-2116

Query Match
Best Local Similarity 100.0%; Score 7; DB 4; Length 55;
Pred. No. 21; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 485 ORAMLEN 491
Db 36 ORAMLEN 42

RESULT 9
US-09-107-433-4073
; Sequence 4073, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4073:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...79
; SEQUENCE DESCRIPTION: SEQ ID NO: 4073:
US-09-107-433-4073

Query Match
Best Local Similarity 100.0%; Score 7; DB 4; Length 79;
Pred. No. 30; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 444 EFKAEKE 450
Db 71 EFKAEKE 77
```

```
RESULT 10
US-09-902-540-15484
; Sequence 15484, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15484
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15484

Query Match      1.4%; Score 7; DB 4; Length 93;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      473 TPVPGI 479      |||||
DB      54 TPVPGI 60

RESULT 11
US-09-270-767-34609
; Sequence 34609, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 34609
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-34609

Query Match      1.4%; Score 7; DB 4; Length 134;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      286 TPVPGI 292      |||||
DB      99 TPVPGI 105

RESULT 12
US-09-270-767-49826
; Sequence 49826, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
```

```
; SEQ ID NO 49826
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-49826

Query Match      1.4%; Score 7; DB 4; Length 134;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      286 TPVPGI 292      |||||
DB      99 TPVPGI 105

RESULT 13
US-09-957-641A-19
; Sequence 19, Application US/09957641A
; Patent No. 6770744
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-00 US
; CURRENT APPLICATION NUMBER: US/09/957,641A
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 60/234,047
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/236,460
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 19
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Porcine
US-09-957-641A-19

Query Match      1.4%; Score 7; DB 4; Length 160;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      476 TPVNAL 482      |||||
DB      119 TPVNAL 125

RESULT 14
US-09-252-991A-24024
; Sequence 24024, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24024
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24024

Query Match      1.4%; Score 7; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 469 APLVPG 475
 DB 85 APLVPG 91

RESULT 15
 US-08-272-255-16

Sequence 16, Application US/08272255
 Patent No. 5824859
 GENERAL INFORMATION:
 APPLICANT: Cashmore, Anthony R.
 APPLICANT: Ahmad, Margaret
 APPLICANT: Lin, Chentao
 TITLE OF INVENTION: Blue light Photoreceptors and Methods of
 TITLE OF INVENTION: Using the same
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859r1s
 STREET: One Liberty Place, 46th floor
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/272,255
 FILING DATE: 08-JUL-1994
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Leary Ph.D., Kathryn
 REGISTRATION NUMBER: 36,317
 REFERENCE/DOCKET NUMBER: UPN-1795
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-3100
 TELEFAX: (215) 568-3439
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 224 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-272-255-16

Query Match 1.4%; Score 7; DB 2; Length 224;
 Best Local Similarity 100.0%; Pred. No. 81;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 421 EDSLAA 427
 DB 8 EDSLAA 14

Search completed: June 7, 2005, 17:11:47
 Job time : 33 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 17:09:41 ; Search time 77.5 Seconds
(without alignments)
2370.369 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 510
Sequence: 1 MFIENFKVCECPVVKYETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 1599520 seqs, 360203123 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1599520

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pap:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pap:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pap:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pap:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pap:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pap:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pap:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pap:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pap:*
17: /cgn2_6/ptodata/1/pubpaa/US10D_NEW_PUB.pap:*
18: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pap:*
19: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pap:*
20: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap:*
21: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	100.0	510	US-10-025-003-2	Sequence 2, Appl1
2	510	100.0	510	US-10-025-003-10	Sequence 10, Appl1
3	510	100.0	510	US-10-718-952-2	Sequence 2, Appl1
4	510	100.0	510	US-10-718-952-10	Sequence 10, Appl1
5	440	86.3	511	US-10-424-599-213009	Sequence 21,009.
6	409	80.2	510	US-10-025-003-6	Sequence 6, Appl1
7	409	80.2	510	US-10-718-952-6	Sequence 6, Appl1
8	253	49.6	510	US-10-025-003-14	Sequence 14, Appl1
9	253	49.6	510	US-10-025-003-16	Sequence 16, Appl1
10	253	49.6	510	US-10-718-952-14	Sequence 14, Appl1
11	253	49.6	510	US-10-718-952-16	Sequence 16, Appl1
12	236	46.3	510	US-10-025-003-12	Sequence 12, Appl1

13	236	46.3	510	US-10-718-952-12	Sequence 12, Appl1
14	159	31.2	431	US-10-424-599-154863	Sequence 154863,
15	113	22.2	220	US-10-424-599-259493	Sequence 259493,
16	103	20.2	211	US-10-424-599-165505	Sequence 165505,
17	102	20.0	510	US-10-424-599-154864	Sequence 154864,
18	73	14.3	250	US-10-424-599-154861	Sequence 154861,
19	68	13.3	332	US-10-425-114-42702	Sequence 42702, A
20	68	13.3	394	US-10-363-829-458	Sequence 458, App
21	68	13.3	450	US-10-425-115-231845	Sequence 231845,
22	68	13.3	510	US-09-727-628-2	Sequence 2, Appl1
23	68	13.3	510	US-10-425-115-231850	Sequence 231850,
24	68	13.3	510	US-10-425-115-231853	Sequence 231853,
25	68	13.3	510	US-10-425-115-231857	Sequence 231857,
26	68	13.3	512	US-10-425-114-39621	Sequence 39621, A
27	68	13.3	512	US-10-425-114-46916	Sequence 46916, A
28	68	13.3	512	US-10-425-114-53231	Sequence 53231, A
29	68	13.3	512	US-10-425-114-66216	Sequence 66216, A
30	67	13.1	512	US-10-425-115-231843	Sequence 231843,
31	67	13.1	505	US-10-442-017-15	Sequence 15, Appl1
32	67	13.1	510	US-09-921-232-11	Sequence 11, Appl1
33	67	13.1	510	US-09-921-330-11	Sequence 11, Appl1
34	67	13.1	510	US-09-921-329-11	Sequence 11, Appl1
35	67	13.1	510	US-10-437-963-173252	Sequence 173252,
36	67	13.1	510	US-10-767-701-46278	Sequence 46278, A
37	67	13.1	510	US-10-425-115-231852	Sequence 231852,
38	67	13.1	510	US-10-425-115-306814	Sequence 306814,
39	67	13.1	510	US-10-739-930-7635	Sequence 7635, Ap
40	67	13.1	516	US-10-425-114-62568	Sequence 62568, A
41	67	13.1	542	US-10-425-114-58674	Sequence 58674, A
42	66	12.9	124	US-10-424-599-213004	Sequence 213004,
43	47	9.2	78	US-10-424-599-213008	Sequence 213008,
44	42	8.2	84	US-10-424-599-154859	Sequence 154859,
45	31	6.1	16	US-10-425-115-231841	Sequence 231841,

ALIGNMENTS

RESULT 1
US-10-025-003-2
; Sequence 2, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strell, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-025-003-2
Query Match 100.0%; Score 510; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFIENFKVCECPVVKYETETIQLVYVYETTELVHNNNGYQWIVVVKYSKYEFKNIHVP 60
DB 1 MFIENFKVCECPVVKYETETIQLVYVYETTELVHNNNGYQWIVVVKYSKYEFKNIHVP 60
QY 61 KLGWMLVGCGNNGSLITGCVIANRBSIWAIRKDKIQANFYGSLQASAIIVGSGQBE 120

```
Db 61 KGVMLVGMGNGNSTLTGGVIANREGISWATKDKIQANVFGSLTQASAIRVGSFGQBE 120
Qy 121 IYAPFKSLIPMNPPDIYFGGWDISNMNLADAMAAKVPDIDLOQLRPYMESMLPLPGI 180
Db 121 IYAPFKSLIPMNPPDIYFGGWDISNMNLADAMAAKVPDIDLOQLRPYMESMLPLPGI 180
Qy 181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIKAFKATKVDKVVVLTANTERTSNLY 240
Db 181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIKAFKATKVDKVVVLTANTERTSNLY 240
Qy 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMENVPFINSPOMTFVPGILDIAIANT 300
Db 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMENVPFINSPOMTFVPGILDIAIANT 300
Qy 301 LIGGDDFKSGGOTKMSVVDPLVAGIKPTSIIVSYNHLGNNDGMNLSAPOTFRSEISKS 360
Db 301 LIGGDDFKSGGOTKMSVVDPLVAGIKPTSIIVSYNHLGNNDGMNLSAPOTFRSEISKS 360
Qy 361 NVVDDMVNSMILYEPGHPDHVVVVKVPPYVGSKRAMEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSMILYEPGHPDHVVVVKVPPYVGSKRAMEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAELSTRIEFKAENEGKFSFHVATILSYLTAKAPLVPPTPVVN 480
Db 421 EDSLAAPIIIDLVLLAELSTRIEFKAENEGKFSFHVATILSYLTAKAPLVPPTPVVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 2

```
US-10-025-003-10
; Sequence 10, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-10
```

```
Query Match 100.0%; Score 510; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MFIENFKVCEPNVKYKTEIEIGSVYNYETTELVEHNRNGTYQWIVPKSVKTEFKTNIHVP 60
Db 1 MFIENFKVCEPNVKYKTEIEIGSVYNYETTELVEHNRNGTYQWIVPKSVKTEFKTNIHVP 60
Qy 61 KGVMLVGMGNGNSTLTGGVIANREGISWATKDKIQANVFGSLTQASAIRVGSFGQBE 120
Db 61 KGVMLVGMGNGNSTLTGGVIANREGISWATKDKIQANVFGSLTQASAIRVGSFGQBE 120
Qy 121 IYAPFKSLIPMNPPDIYFGGWDISNMNLADAMAAKVPDIDLOQLRPYMESMLPLPGI 180
Db 121 IYAPFKSLIPMNPPDIYFGGWDISNMNLADAMAAKVPDIDLOQLRPYMESMLPLPGI 180
Qy 181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIKAFKATKVDKVVVLTANTERTSNLY 240
Db 181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIKAFKATKVDKVVVLTANTERTSNLY 240
```

```
Qy 181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIKAFKATKVDKVVVLTANTERTSNLY 240
Db 181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIKAFKATKVDKVVVLTANTERTSNLY 240
Qy 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMENVPFINSPOMTFVPGILDIAIANT 300
Db 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMENVPFINSPOMTFVPGILDIAIANT 300
Qy 301 LIGGDDFKSGGOTKMSVVDPLVAGIKPTSIIVSYNHLGNNDGMNLSAPOTFRSEISKS 360
Db 301 LIGGDDFKSGGOTKMSVVDPLVAGIKPTSIIVSYNHLGNNDGMNLSAPOTFRSEISKS 360
Qy 361 NVVDDMVNSMILYEPGHPDHVVVVKVPPYVGSKRAMEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSMILYEPGHPDHVVVVKVPPYVGSKRAMEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAELSTRIEFKAENEGKFSFHVATILSYLTAKAPLVPPTPVVN 480
Db 421 EDSLAAPIIIDLVLLAELSTRIEFKAENEGKFSFHVATILSYLTAKAPLVPPTPVVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 3

```
US-10-718-952-2
; Sequence 2, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-2
```

```
Query Match 100.0%; Score 510; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MFIENFKVCEPNVKYKTEIEIGSVYNYETTELVEHNRNGTYQWIVPKSVKTEFKTNIHVP 60
Db 1 MFIENFKVCEPNVKYKTEIEIGSVYNYETTELVEHNRNGTYQWIVPKSVKTEFKTNIHVP 60
Qy 61 KGVMLVGMGNGNSTLTGGVIANREGISWATKDKIQANVFGSLTQASAIRVGSFGQBE 120
Db 61 KGVMLVGMGNGNSTLTGGVIANREGISWATKDKIQANVFGSLTQASAIRVGSFGQBE 120
Qy 121 IYAPFKSLIPMNPPDIYFGGWDISNMNLADAMAAKVPDIDLOQLRPYMESMLPLPGI 180
Db 121 IYAPFKSLIPMNPPDIYFGGWDISNMNLADAMAAKVPDIDLOQLRPYMESMLPLPGI 180
Qy 181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIKAFKATKVDKVVVLTANTERTSNLY 240
Db 181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIKAFKATKVDKVVVLTANTERTSNLY 240
Qy 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMENVPFINSPOMTFVPGILDIAIANT 300
```

```
Db 241 VGLNDTMENTLAAVDNREAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIRNT 300
Qy 301 LIGGDDPKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLAPQTFRSKEISKS 360
Db 301 LIGGDDPKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLAPQTFRSKEISKS 360
Qy 361 NVVDDMVNSNALIYERGEHPDHVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYERGEHPDHVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAEISTRIEFAKNEGKFHSPVATILSYLTAKAPLVPGTPPVN 480
Db 421 EDSLAAPIIIDLVLLAEISTRIEFAKNEGKFHSPVATILSYLTAKAPLVPGTPPVN 480
Qy 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
```

RESULT 4

```
US-10-718-952-10
; Sequence 10, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strelt, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-10
```

```
Query Match 100.0%; Score 510; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MEIENFKVECPNVKTYTETEIQSYVNYETTELVEHNRNGTYQWIVKPKSVKYEFTNIHVP 60
Db 1 MEIENFKVECPNVKTYTETEIQSYVNYETTELVEHNRNGTYQWIVKPKSVKYEFTNIHVP 60
Qy 61 KGVNMLVGKGNNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGSEE 120
Db 61 KGVNMLVGKGNNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGSEE 120
Qy 121 IYAPFKSLLPVNPPDIVFGWDISNMNLADAMARAKVFDIDLQKOLRPYMSMLPLPGI 180
Db 121 IYAPFKSLLPVNPPDIVFGWDISNMNLADAMARAKVFDIDLQKOLRPYMSMLPLPGI 180
Qy 181 YPDPFLAANOEBRANNVIGTKOEOVOQIIKDKAKFEATKYDKVVVLTANTERYSNLV 240
Db 181 YPDPFLAANOEBRANNVIGTKOEOVOQIIKDKAKFEATKYDKVVVLTANTERYSNLV 240
Qy 241 VGLNDTMENTLAAVDNREAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIRNT 300
Db 241 VGLNDTMENTLAAVDNREAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIRNT 300
Qy 301 LIGGDDPKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLAPQTFRSKEISKS 360
Db 301 LIGGDDPKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLAPQTFRSKEISKS 360
```

```
Qy 361 NVVDDMVNSNALIYERGEHPDHVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYERGEHPDHVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAEISTRIEFAKNEGKFHSPVATILSYLTAKAPLVPGTPPVN 480
Db 421 EDSLAAPIIIDLVLLAEISTRIEFAKNEGKFHSPVATILSYLTAKAPLVPGTPPVN 480
Qy 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
```

RESULT 5

```
US-10-424-599-213009
; Sequence 213009, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 213009
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(511)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1.pep
US-10-424-599-213009
```

```
Query Match 86.3%; Score 440; DB 15; Length 511;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 71 GNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGSEEIYAPFKSLLP 130
Db 71 GNGSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGSEEIYAPFKSLLP 130
Qy 131 MNPPDDIYFGWDISNMNLADAMARAKVFDIDLQKOLRPYMSMLPLPGIYPPDFIAAQ 190
Db 131 MNPPDDIYFGWDISNMNLADAMARAKVFDIDLQKOLRPYMSMLPLPGIYPPDFIAAQ 190
Qy 191 EERANNVIGTKOEOVOQIIKDKAKFEATKYDKVVVLTANTERYSNLVGLNDTMENT 250
Db 191 EERANNVIGTKOEOVOQIIKDKAKFEATKYDKVVVLTANTERYSNLVGLNDTMENT 250
Qy 251 LAAVDNREAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIRNTLIGGDDPKSG 310
Db 251 LAAVDNREAEISPTLYALACWENVPFINGSPQNTFVGLIDLAIRNTLIGGDDPKSG 310
Qy 311 QTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLAPQTFRSKEISKNVDDMVNSN 370
Db 311 QTKMSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLAPQTFRSKEISKNVDDMVNSN 370
Qy 371 ALIYEPGEHPDHVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTCESLLAAPII 430
Db 371 ALIYEPGEHPDHVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTCESLLAAPII 430
Qy 431 IDLVLLAEISTRIEFAKNEGKFHSPVATILSYLTAKAPLVPGTPPVNALS KORAMLE 490
Db 431 IDLVLLAEISTRIEFAKNEGKFHSPVATILSYLTAKAPLVPGTPPVNALS KORAMLE 490
```

Qy	491	NIMRACVGLAPENNMILEYK	510
Db	492	NIMRACVGLAPENNMILEYK	511

```

RESULT 6
US-10-025-003-6
Sequence 6, Application US/10025003
Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streif, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 6
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-025-003-6

```

Query Match	80.2%	Score 409;	DB 14;	Length 510;
Best Local Similarity	99.8%	Pred. No. 0;		
Matches 509;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	MFIEFNKVECHGVKXKTELETETIOSVYNYVETTELVEHNRNGTGYMIVKPKSVKTEFKTNHVP	60	
Db	1	MFIEFNKVECHGVKXKTELETETIOSVYNYVETTELVEHNRNGTGYMIVKPKSVKTEFKTNHVP	60	
Qy	61	KLGMVLVGMGNGNSTLTGVIYANREGISWATPKDIOQANFYGLTQASAIRVGSFGQEE	120	
Db	61	KLGMVLVGMGNGNSTLTGVIYANREGISWATPKDIOQANFYGLTQASAIRVGSFGQEE	120	
Qy	121	IYAPFKSLLEPVNPDIDIVFGGWDISNNMLADAMARAKVPDIDLOKOLRPYMESMLPLPGI	180	
Db	121	IYAPFKSLLEPVNPDIDIVFGGWDISNNMLADAMARAKVPDIDLOKOLRPYMESMLPLPGI	180	
Qy	181	YDPDFIANOBERRANNVAKGTROBOVOOIIKDIAFKRATYVDKVVYLMTANTEYSMLV	240	
Db	181	YDPDFIANOBERRANNVAKGTROBOVOOIKDIAFKRATYVDKVVYLMTANTEYSMLV	240	
Qy	241	VGNDTMENTLLAAVDNRNEAISPSTLYIAACMENVPFINGSPONTFVPGILDILAIAANT	300	
Db	241	VGNDTMENTLLAAVDNRNEAISPSTLYIAACMENVPFINGSPONTFVPGILDILAIAANT	300	
Qy	301	LIIGDDFESGQTKMSVYVDFLVGAGIPTSTIVSNHIGNDGMLSPQTFRSKEISKS	360	
Db	301	LIIGDDFESGQTKMSVYVDFLVGAGIPTSTIVSNHIGNDGMLSPQTFRSKEISKS	360	
Qy	361	NVVDMDVNSNAIIVERGEHPDHVVVIAKVVPYVYDGSKRAMDEYTSEIFPGKSTIVLHNTC	420	
Db	361	NVVDMDVNSNAIIVERGEHPDHVVVIAKVVPYVYDGSKRAMDEYTSEIFPGKSTIVLHNTC	420	
Qy	421	EDSLIAAIIIDVLVLAELSTRIEFKANEGCFHSFHVATILSLYTKAPLVLPBGTPIVNN	480	
Db	421	EDSLIAAIIIDVLVLAELSTRIEFKANEGCFHSFHVATILSLYTKAPLVLPBGTPIVNN	480	
Qy	481	ALSKORAMLENIMRACVGLAPENNMILEYK	510	
Db	481	ALSKORAMLENIMRACVGLAPENNMILEYK	510	

```

US-10-718-952-6
; Sequence 6, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-6

Query Match      80.2%; Score 409; DB 16; Length 510;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Query Match	80.2%	Score 409	DB 16	Length 510
Best Local Similarity	99.8%	Pred. No. 0		
Matches 509	Conservative 0	Mismatches 1	Indels 0	Gaps 0

Qy	1	MFENFKVECBNVKXTEIETIQSVYNYETTELVHENRNGTQYIVKPKSVKTEFKTNHVP	60
Db	1	MFENFKVECBNVKXTEIETIQSVYNYETTELVHENRNGTQYIVKPKSVKTEFKTNHVP	60
Qy	61	KLGMVLWGMGNNNSTLGTGVIARREGISMAKTHXKIQOANYFGSLTQSAIRVSGFQGBE	120
Db	61	KLGMVLWGMGNNNSTLGTGVIARREGISMAKTHXKIQOANYFGSLTQSAIRVSGFQGBE	120
Qy	121	IYAFPKSLPWNPDIDIVFGMDISNNMLADAMARAKFIDIDLOKOLRPYMESMLPLPGI	180
Db	121	IYAFPKSLPWNPDIDIVFGMDISNNMLADAMARAKFIDIDLOKOLRPYMESMLPLPGI	180
Qy	181	YDPDFPIANQBERANNVVKGTQEOVOQIIDIKAFKATRYDKVYVMTANTERYSNLV	240
Db	181	YDPDFPIANQBERANNVVKGTQEOVOQIIDIKAFKATRYDKVYVMTANTERYSNLV	240
Qy	241	VGUNDMENTLAAVDRNBAEISPTSLVIAACMVENVPTINGSPOVTFVGLIDLIAART	300
Db	241	VGUNDMENTLAAVDRNBAEISPTSLVIAACMVENVPTINGSPOVTFVGLIDLIAART	300
Qy	301	LIGGDFFKSGOTKKKSVLVDFLVGAGIKPTISIVSYNHLGNNDMNLSAPQTFRSKEISKS	360
Db	301	LIGGDFFKSGOTKKKSVLVDFLVGAGIKPTISIVSYNHLGNNDMNLSAPQTFRSKEISKS	360
Qy	361	NVVDMDNNSNALIYEPGHPHHVYIKVPPYVGSKRAMDEYTSIIFMGKSTIYLAHNTC	420
Db	361	NVVDMDNNSNALIYEPGHPHHVYIKVPPYVGSNRMADYTSIIFMGKSTIYLAHNTC	420
Qy	421	EDSLLAAPIIDLVLAELSTRIEFKANEKGKFSFHHVATILSLYTAAPLVPGTTPVYN	480
Db	421	EDSLLAAPIIDLVLAELSTRIEFKANEKGKFSFHHVATILSLYTAAPLVPGTTPVYN	480
Qy	481	ALSKORAMLENIMRACVGLAEPENNMILEYK 510	
Db	481	ALSKORAMLENIMRACVGLAEPENNMILEYK 510	

RESULT 8
US-10-025-003-14
; Sequence 14, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott

APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/935,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 14
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-025-003-14

Query Match 49.6%; Score 253; DB 14; Length 510;
Best Local Similarity 99.7%; Pred. No. 7.9e-242;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 HYPKGLVNLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFSGSLTQASAIRVGSFQ 117
DB 58 HYPKGLVNLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFSGSLTQASAIRVGSFQ 117
QY 118 GEEIYAPFKSLLPWNPDDIVFGGWDISNMNLADAMARAKVPDIDLOKOLRPYMSMLPL 177
DB 118 GEEIYAPFKSLLPWNPDDIVFGGWDISNMNLADAMARAKVPDIDLOKOLRPYMSMLPL 177
QY 178 PGTYDPDFIAANOEBERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVWLTANTERS 237
DB 178 PGTYDPDFIAANOEBERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVWLTANTERS 237
QY 238 NLVGLNDTMENTLLAAVDNBEAISPSTLYALACWENVPFINGSQNTFVPGILDALA 297
DB 238 NLVGLNDTMENTLLAAVDNBEAISPSTLYALACWENVPFINGSQNTFVPGILDALA 297
QY 298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKPTSIYSYNHLGNNDGMNLSAPQTFRSKEI 357
DB 298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKPTSIYSYNHLGNNDGMNLSAPQTFRSKEI 357
QY 358 SXSNNVDDMVNSNALIYEPGEHPDHVVYIKYVPYVDSGRANDEYTSSEIFMGK 411
DB 358 SXSNNVDDMVNSNALIYEPGEHPDHVVYIKYVPYVDSGRANDEYTSSEIFMGK 411

RESULT 9

US-10-025-003-16
Sequence 16, Application US/10025003
Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/935,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 16
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-025-003-16

Query Match 49.6%; Score 253; DB 14; Length 510;
Best Local Similarity 99.7%; Pred. No. 7.9e-242;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 HYPKGLVNLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFSGSLTQASAIRVGSFQ 117
DB 58 HYPKGLVNLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFSGSLTQASAIRVGSFQ 117
QY 118 GEEIYAPFKSLLPWNPDDIVFGGWDISNMNLADAMARAKVPDIDLOKOLRPYMSMLPL 177
DB 118 GEEIYAPFKSLLPWNPDDIVFGGWDISNMNLADAMARAKVPDIDLOKOLRPYMSMLPL 177
QY 178 PGTYDPDFIAANOEBERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVWLTANTERS 237
DB 178 PGTYDPDFIAANOEBERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVWLTANTERS 237
QY 238 NLVGLNDTMENTLLAAVDNBEAISPSTLYALACWENVPFINGSQNTFVPGILDALA 297
DB 238 NLVGLNDTMENTLLAAVDNBEAISPSTLYALACWENVPFINGSQNTFVPGILDALA 297
QY 298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKPTSIYSYNHLGNNDGMNLSAPQTFRSKEI 357
DB 298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKPTSIYSYNHLGNNDGMNLSAPQTFRSKEI 357
QY 358 SXSNNVDDMVNSNALIYEPGEHPDHVVYIKYVPYVDSGRANDEYTSSEIFMGK 411
DB 358 SXSNNVDDMVNSNALIYEPGEHPDHVVYIKYVPYVDSGRANDEYTSSEIFMGK 411

RESULT 10

US-10-718-952-14
Sequence 14, Application US/10718952
Publication No. US20040128713A1
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/718,952
CURRENT FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: 08/935,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 14
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-718-952-14

Query Match 49.6%; Score 253; DB 16; Length 510;
Best Local Similarity 99.7%; Pred. No. 7.9e-242;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 HYPKGLVNLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFSGSLTQASAIRVGSFQ 117
DB 58 HYPKGLVNLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFSGSLTQASAIRVGSFQ 117
QY 118 GEEIYAPFKSLLPWNPDDIVFGGWDISNMNLADAMARAKVPDIDLOKOLRPYMSMLPL 177
DB 118 GEEIYAPFKSLLPWNPDDIVFGGWDISNMNLADAMARAKVPDIDLOKOLRPYMSMLPL 177
QY 178 PGTYDPDFIAANOEBERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVWLTANTERS 237
DB 178 PGTYDPDFIAANOEBERANNVIGTKQEOVOQIIKDIKAFKATKVDKVVWLTANTERS 237
QY 238 NLVGLNDTMENTLLAAVDNBEAISPSTLYALACWENVPFINGSQNTFVPGILDALA 297

```

Db      238 NLVVGINDTMENTLAADVDRNEAISPSTLYALACWENVPFINGSPONTFVGLIDLALIA 297
      298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKPTSTISVYNHLGNNDGMNLSAPQTFRSKEI 357
Db      298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKPTSTISVYNHLGNNDGMNLSAPQTFRSKEI 357
Qy      358 SKSNVVDVDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
      358 SKSNVVDVDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
Db      358 SKSNVVDVDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411

```

RESULT 11

```

US-10-718-952-16
; Sequence 16, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-16

```

```

Query Match      49.6%; Score 253; DB 16; Length 510;
Best Local Similarity 99.7%; Pred. No. 7, 9e-242;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      58 HVPKLGVMVLVGNGNGSTLTGGVIANREGISWATKDKIQANYPGSLTQASAIRVGSFO 117
      58 HVPKLGVMVLVGNGNGSTLTGGVIANREGISWATKDKIQANYPGSLTQASAIRVGSFO 117
Db      118 GEIYAPFESLPMVNPDDIVFGMDISNMNLADAAKAKVFDIDLOKOLRPFMESMBL 177
      118 GEIYAPFESLPMVNPDDIVFGMDISNMNLADAAKAKVFDIDLOKOLRPFMESMBL 177
Db      178 PGYIDPDFIAANOEBEERANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERYS 237
      178 PGYIDPDFIAANOEBEERANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERYS 237
Qy      238 NLVVGINDTMENTLAADVDRNEAISPSTLYALACWENVPFINGSPONTFVGLIDLALIA 297
      238 NLVVGINDTMENTLAADVDRNEAISPSTLYALACWENVPFINGSPONTFVGLIDLALIA 297
Db      238 NLVVGINDTMENTLAADVDRNEAISPSTLYALACWENVPFINGSPONTFVGLIDLALIA 297
Qy      298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKPTSTISVYNHLGNNDGMNLSAPQTFRSKEI 357
      298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKPTSTISVYNHLGNNDGMNLSAPQTFRSKEI 357
Db      298 RNTLIGDDPKSGQTKMSVLVDPLVAGIKPTSTISVYNHLGNNDGMNLSAPQTFRSKEI 357
Qy      358 SKSNVVDVDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
      358 SKSNVVDVDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
Db      358 SKSNVVDVDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411

```

RESULT 12

```

US-10-025-003-12
; Sequence 12, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William

```

```

; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-12

```

```

Query Match      46.3%; Score 236; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 5, 8e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      176 PLPGYIDPDFIAANOEBEERANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTER 235
      176 PLPGYIDPDFIAANOEBEERANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTER 235
Db      236 YSNLVGINDTMENTLAADVDRNEAISPSTLYALACWENVPFINGSPONTFVGLIDLALIA 295
      236 YSNLVGINDTMENTLAADVDRNEAISPSTLYALACWENVPFINGSPONTFVGLIDLALIA 295
Qy      296 IARNTLIGDDPKSGQTKMSVLVDPLVAGIKPTSTISVYNHLGNNDGMNLSAPQTFRSK 355
      296 IARNTLIGDDPKSGQTKMSVLVDPLVAGIKPTSTISVYNHLGNNDGMNLSAPQTFRSK 355
Db      356 EISKSNVVDVDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411
      356 EISKSNVVDVDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSCRANDEYTSSEIFMGSK 411

```

RESULT 13

```

US-10-718-952-12
; Sequence 12, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-12

```

```

Query Match      46.3%; Score 236; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 5, 8e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      176 PLPGYIDPDFIAANOEBEERANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTER 235
      176 PLPGYIDPDFIAANOEBEERANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTER 235

```

```
Db      176 PLPGIYDPDFIAANOEBRANNVIKGTKEQVOQI1KIDIAKFAKATKYDKVVLMTANTER 235
Qy      236 YSNLVVGLNDTMENTNLAADVDRNEAISPSTLYAIACMENVPFINGSPONTFVPGIIDL 295
Db      236 YSNLVVGLNDTMENTNLAADVDRNEAISPSTLYAIACMENVPFINGSPONTFVPGIIDL 295
Qy      296 IARNTLIGDDPFSGQTKKMSVLDVFLVGAGIKPTISIVSYNHLGNNDGMLSAPOTFRSK 355
Db      296 IARNTLIGDDPFSGQTKKMSVLDVFLVGAGIKPTISIVSYNHLGNNDGMLSAPOTFRSK 355
Qy      356 EISKSNVVDMDVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMEYTSIEIFMGSK 411
Db      356 EISKSNVVDMDVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMEYTSIEIFMGSK 411
```

```
RESULT 14
US-10-424-599-154863
; Sequence 154863, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154863
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110862C.1.pep
US-10-424-599-154863
```

```
Query Match      31.2%; Score 159; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.2e-148;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      236 YSNLVVGLNDTMENTNLAADVDRNEAISPSTLYAIACMENVPFINGSPONTFVPGIIDL 295
Db      236 YSNLVVGLNDTMENTNLAADVDRNEAISPSTLYAIACMENVPFINGSPONTFVPGIIDL 295
Qy      296 IARNTLIGDDPFSGQTKKMSVLDVFLVGAGIKPTISIVSYNHLGNNDGMLSAPOTFRSK 355
Db      296 IARNTLIGDDPFSGQTKKMSVLDVFLVGAGIKPTISIVSYNHLGNNDGMLSAPOTFRSK 355
Qy      356 EISKSNVVDMDVNSNALIYEPGEHPDHVVVIKYVPYVGD 394
Db      356 EISKSNVVDMDVNSNALIYEPGEHPDHVVVIKYVPYVGD 394
```

```
RESULT 15
US-10-424-599-259439
; Sequence 259439, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 259439
; LENGTH: 220
; TYPE: PRT
```

```
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(220)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_76298C.1.pep
US-10-424-599-259439
```

```
Query Match      22.2%; Score 113; DB 15; Length 220;
Best Local Similarity 100.0%; Pred. No. 2.9e-103;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      256 RNEAISPSTLYAIACMENVPFINGSPONTFVPGIIDLIAIARNTLIGDDPFSGQTKK 315
Db      2 RNEAISPSTLYAIACMENVPFINGSPONTFVPGIIDLIAIARNTLIGDDPFSGQTKK 61
Qy      316 SYLVDFLVGAGIKPTISIVSYNHLGNNDGMLSAPOTFRSKEISKSNVVDMDVN 368
Db      62 SYLVDFLVGAGIKPTISIVSYNHLGNNDGMLSAPOTFRSKEISKSNVVDMDVN 114
```

```
Search completed: June 7, 2005, 17:25:33
Job time : 78.5 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 17:03:16 ; Search time 276.5 Seconds
(without alignments)
2154.370 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 510
Sequence: 1 MEIENPKVCSPVKYETETI.....NIMRACVGLAPNNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 6959266 seqs, 116806243 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Pending Patents AA Main:*

1: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
33: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
34: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
35: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
36: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
37: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	510	100.0	510	12	US-08-835-751A-2	Sequence 2, Appl1
2	510	100.0	510	16	US-09-299-315-2	Sequence 2, Appl1
3	510	100.0	510	16	US-09-299-315-10	Sequence 10, Appl1
4	510	100.0	510	17	US-09-304-534-2	Sequence 2, Appl1
5	510	100.0	510	17	US-09-304-534-10	Sequence 10, Appl1
6	510	100.0	510	26	US-10-025-003-2	Sequence 2, Appl1
7	510	100.0	510	26	US-10-025-003-10	Sequence 10, Appl1
8	510	100.0	510	33	US-10-718-952-2	Sequence 2, Appl1
9	510	100.0	510	33	US-10-718-952-10	Sequence 10, Appl1
10	510	100.0	510	37	US-60-592-978-4180	Sequence 4180, Ap
11	510	100.0	510	37	US-60-592-978-6132	Sequence 6132, Ap
12	440	86.3	511	30	US-10-424-559-213009	Sequence 213009,
13	409	80.2	510	12	US-08-835-751A-6	Sequence 6, Appl1
14	409	80.2	510	16	US-09-299-315-6	Sequence 6, Appl1
15	409	80.2	510	17	US-09-304-534-6	Sequence 6, Appl1
16	409	80.2	510	26	US-10-025-003-6	Sequence 6, Appl1
17	409	80.2	510	33	US-10-718-952-6	Sequence 6, Appl1
18	376	73.7	510	37	US-60-592-978-13648	Sequence 13648, A
19	253	49.6	510	16	US-09-299-315-14	Sequence 16, Appl1
20	253	49.6	510	16	US-09-299-315-16	Sequence 16, Appl1
21	253	49.6	510	17	US-09-304-534-14	Sequence 14, Appl1
22	253	49.6	510	17	US-09-304-534-16	Sequence 16, Appl1
23	253	49.6	510	26	US-10-025-003-14	Sequence 14, Appl1
24	253	49.6	510	26	US-10-025-003-16	Sequence 16, Appl1
25	253	49.6	510	33	US-10-718-952-14	Sequence 14, Appl1
26	253	49.6	510	33	US-10-718-952-16	Sequence 16, Appl1
27	253	49.6	510	37	US-60-592-978-12850	Sequence 12850, A
28	236	46.3	510	17	US-09-299-315-12	Sequence 12, Appl1
29	236	46.3	510	16	US-09-304-534-12	Sequence 12, Appl1
30	236	46.3	510	26	US-10-025-003-12	Sequence 12, Appl1
31	236	46.3	510	33	US-10-718-952-12	Sequence 12, Appl1
32	165	32.4	636	37	US-60-592-978-15981	Sequence 15981, A
33	159	31.2	431	30	US-10-424-559-15483	Sequence 15483,
34	115	22.5	184	37	US-60-592-978-2477	Sequence 2477, Ap
35	113	22.2	220	30	US-10-424-559-259439	Sequence 259439,
36	113	22.2	220	37	US-60-592-978-10124	Sequence 10124, A
37	111	21.8	260	37	US-60-592-978-20424	Sequence 20424, A
38	110	21.6	148	37	US-60-592-978-7377	Sequence 7377, Ap
39	103	20.2	211	30	US-10-424-559-165505	Sequence 165505,
40	103	20.2	276	37	US-60-592-978-21322	Sequence 21322, A
41	102	20.0	510	30	US-10-424-559-154864	Sequence 154864,
42	97	19.0	211	37	US-60-592-978-4373	Sequence 4373, Ap
43	94	18.4	557	37	US-60-592-978-21798	Sequence 21798, A
44	92	18.0	472	37	US-60-592-978-4518	Sequence 4518, Ap
45	79	15.5	226	37	US-60-592-978-15017	Sequence 15017, A

ALIGNMENTS

RESULT 1
US-08-835-751A-2
; Sequence 2, Application US/08835751A
; GENERAL INFORMATION:
; APPLICANT: HITZ, WILLIAM D.
; APPLICANT: SEBASTIAN, SCOTT ANTHONY
; TITLE OF INVENTION: SOYBEAN PLANTS PRODUCING SEEDS WITH
; TITLE OF INVENTION: REDUCED LEVELS OF RAFFINOSE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH

COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95 (7.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,751A
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: MAJARIAN, WILLIAM R.
REGISTRATION NUMBER: P41,173
REFERENCE/DOCKET NUMBER: BB-1077
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302)992-4926
TELEFAX: (302)773-0164
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-835-751A-2

Query Match 100.0%; Score 510; DB 12; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLENFKVCEPNVKTETEIQSVYNYETTELVEHNRNGTYOMIVKPKSVKTEFKTNHVP 60
DB 1 MFLENFKVCEPNVKTETEIQSVYNYETTELVEHNRNGTYOMIVKPKSVKTEFKTNHVP 60
QY 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDKIQOANTFGSLTQASAIRVGSFOGEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDKIQOANTFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPWNPDPDIVFGGMDISNMNLADAMARAKVFDIDLOKQLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPWNPDPDIVFGGMDISNMNLADAMARAKVFDIDLOKQLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIKAFKATKVDKVVLTMTANTERSMLV 240
DB 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIKAFKATKVDKVVLTMTANTERSMLV 240
QY 241 VGLNDTMENLLAADRNEAISPSTLYAIACMENVPFINGSPOTFVGLIDLAIRNT 300
DB 241 VGLNDTMENLLAADRNEAISPSTLYAIACMENVPFINGSPOTFVGLIDLAIRNT 300
QY 301 LIIGDDFKSGQTKMKSVLVDPLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
DB 301 LIIGDDFKSGQTKMKSVLVDPLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
QY 361 NVYDDMVNSNALIYERGGHPDHVVYIKVYPVYGDSCRAMDEYTSSEIFMGKSTIYLHNTC 420
DB 361 NVYDDMVNSNALIYERGGHPDHVVYIKVYPVYGDSCRAMDEYTSSEIFMGKSTIYLHNTC 420
QY 421 EDSLAAPIIIDLVLLAEISTRIEFKAENEGKFSFHVATILSYLTAKPLVPGTPPVN 480
DB 421 EDSLAAPIIIDLVLLAEISTRIEFKAENEGKFSFHVATILSYLTAKPLVPGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 2
US-09-299-315-2
Sequence 2, Application US/09299315
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streit, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID

FILE REFERENCE: BB-1077-B
CURRENT APPLICATION NUMBER: US/09/299,315
CURRENT FILING DATE: 1999-04-26
EARLIER APPLICATION NUMBER: 08/835,751
EARLIER FILING DATE: APRIL 8, 1997
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 2
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-09-299-315-2

Query Match 100.0%; Score 510; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLENFKVCEPNVKTETEIQSVYNYETTELVEHNRNGTYOMIVKPKSVKTEFKTNHVP 60
DB 1 MFLENFKVCEPNVKTETEIQSVYNYETTELVEHNRNGTYOMIVKPKSVKTEFKTNHVP 60
QY 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDKIQOANTFGSLTQASAIRVGSFOGEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDKIQOANTFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPWNPDPDIVFGGMDISNMNLADAMARAKVFDIDLOKQLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPWNPDPDIVFGGMDISNMNLADAMARAKVFDIDLOKQLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIKAFKATKVDKVVLTMTANTERSMLV 240
DB 181 YDPDFIAANOEBRANNVKGTKEOVQOIIKDIKAFKATKVDKVVLTMTANTERSMLV 240
QY 241 VGLNDTMENLLAADRNEAISPSTLYAIACMENVPFINGSPOTFVGLIDLAIRNT 300
DB 241 VGLNDTMENLLAADRNEAISPSTLYAIACMENVPFINGSPOTFVGLIDLAIRNT 300
QY 301 LIIGDDFKSGQTKMKSVLVDPLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
DB 301 LIIGDDFKSGQTKMKSVLVDPLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
QY 361 NVYDDMVNSNALIYERGGHPDHVVYIKVYPVYGDSCRAMDEYTSSEIFMGKSTIYLHNTC 420
DB 361 NVYDDMVNSNALIYERGGHPDHVVYIKVYPVYGDSCRAMDEYTSSEIFMGKSTIYLHNTC 420
QY 421 EDSLAAPIIIDLVLLAEISTRIEFKAENEGKFSFHVATILSYLTAKPLVPGTPPVN 480
DB 421 EDSLAAPIIIDLVLLAEISTRIEFKAENEGKFSFHVATILSYLTAKPLVPGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 3
US-09-299-315-10
Sequence 10, Application US/09299315
GENERAL INFORMATION:
APPLICANT: Hiltz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streit, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
FILE REFERENCE: BB-1077-B
CURRENT APPLICATION NUMBER: US/09/299,315
CURRENT FILING DATE: 1999-04-26
EARLIER APPLICATION NUMBER: 08/835,751
EARLIER FILING DATE: APRIL 8, 1997
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 10
LENGTH: 510

```

; TYPE: PRt
; ORGANISM: Glycine max
US-09-299-315-10

Query Match
Best Local Similarity 100.0%; Score 510; DB 16; Length 510;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLENKVCPCNVKKTETETIQSVNYVYETTELVEHNRNGTYQWIVKPKSVKYEKTNIHVP 60
DB 1 MFLENKVCPCNVKKTETETIQSVNYVYETTELVEHNRNGTYQWIVKPKSVKYEKTNIHVP 60
QY 61 KGVMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGGE 120
DB 61 KGVMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGGE 120
QY 121 IYAPFKSLIPMVPDDIVFGWMDISNMNLADAMARAKVPDIDLQKLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPMVPDDIVFGWMDISNMNLADAMARAKVPDIDLQKLRPYMESMLPLPGI 180
QY 181 YPDPFIAANOEBRANNVIGTKOEOVQOIIKDIKAFKATKVDKVVVLTANTERSNLV 240
DB 181 YPDPFIAANOEBRANNVIGTKOEOVQOIIKDIKAFKATKVDKVVVLTANTERSNLV 240
QY 241 VGLNDTMENLAAVDNRBAISPSTLYAICVMEVNPFIKNGSPONTFVGLIDLAIARNT 300
DB 241 VGLNDTMENLAAVDNRBAISPSTLYAICVMEVNPFIKNGSPONTFVGLIDLAIARNT 300
QY 301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNNDGMNLSAPQTFRSKEISKS 360
DB 301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNNDGMNLSAPQTFRSKEISKS 360
QY 361 NVYDDMVNSNALIYEBGEHPDHVVVIKYVPYVGDSCRADDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVYDDMVNSNALIYEBGEHPDHVVVIKYVPYVGDSCRADDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPIIIDVLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKPLVPPGTPVVA 480
DB 421 EDSLLAAPIIIDVLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKPLVPPGTPVVA 480
QY 481 ALSKORAMLENIIRACVGLAPENNMLEYK 510
DB 481 ALSKORAMLENIIRACVGLAPENNMLEYK 510

RESULT 4
US-09-304-534-2
; Sequence 2, Application US/09304534
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebaastian, Scott
; APPLICANT: Grace, John
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/09/304, 534
; EARLIER FILING DATE: 1999-05-04
; EARLIER APPLICATION NUMBER: 08/835, 751
; EARLIER FILING DATE: APRIL 8, 1997
; EARLIER APPLICATION NUMBER: PCT/US98/06822
; EARLIER FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRt
; ORGANISM: Glycine max
US-09-304-534-2

Query Match
Best Local Similarity 100.0%; Score 510; DB 17; Length 510;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 MFLENKVCPCNVKKTETETIQSVNYVYETTELVEHNRNGTYQWIVKPKSVKYEKTNIHVP 60
DB 1 MFLENKVCPCNVKKTETETIQSVNYVYETTELVEHNRNGTYQWIVKPKSVKYEKTNIHVP 60
QY 61 KGVMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGGE 120
DB 61 KGVMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGGE 120
QY 121 IYAPFKSLIPMVPDDIVFGWMDISNMNLADAMARAKVPDIDLQKLRPYMESMLPLPGI 180
DB 121 IYAPFKSLIPMVPDDIVFGWMDISNMNLADAMARAKVPDIDLQKLRPYMESMLPLPGI 180
QY 181 YPDPFIAANOEBRANNVIGTKOEOVQOIIKDIKAFKATKVDKVVVLTANTERSNLV 240
DB 181 YPDPFIAANOEBRANNVIGTKOEOVQOIIKDIKAFKATKVDKVVVLTANTERSNLV 240
QY 241 VGLNDTMENLAAVDNRBAISPSTLYAICVMEVNPFIKNGSPONTFVGLIDLAIARNT 300
DB 241 VGLNDTMENLAAVDNRBAISPSTLYAICVMEVNPFIKNGSPONTFVGLIDLAIARNT 300
QY 301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNNDGMNLSAPQTFRSKEISKS 360
DB 301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIYSYNHLGNNDGMNLSAPQTFRSKEISKS 360
QY 361 NVYDDMVNSNALIYEBGEHPDHVVVIKYVPYVGDSCRADDEYTSSEIFMGKSTIVLHNTC 420
DB 361 NVYDDMVNSNALIYEBGEHPDHVVVIKYVPYVGDSCRADDEYTSSEIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPIIIDVLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKPLVPPGTPVVA 480
DB 421 EDSLLAAPIIIDVLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKPLVPPGTPVVA 480
QY 481 ALSKORAMLENIIRACVGLAPENNMLEYK 510
DB 481 ALSKORAMLENIIRACVGLAPENNMLEYK 510

RESULT 5
US-09-304-534-10
; Sequence 10, Application US/09304534
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebaastian, Scott
; APPLICANT: Grace, John
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/09/304, 534
; EARLIER FILING DATE: 1999-05-04
; EARLIER APPLICATION NUMBER: 08/835, 751
; EARLIER FILING DATE: APRIL 8, 1997
; EARLIER APPLICATION NUMBER: PCT/US98/06822
; EARLIER FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRt
; ORGANISM: Glycine max
US-09-304-534-10

Query Match
Best Local Similarity 100.0%; Score 510; DB 17; Length 510;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLENKVCPCNVKKTETETIQSVNYVYETTELVEHNRNGTYQWIVKPKSVKYEKTNIHVP 60
DB 1 MFLENKVCPCNVKKTETETIQSVNYVYETTELVEHNRNGTYQWIVKPKSVKYEKTNIHVP 60
QY 61 KGVMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGGE 120
DB 61 KGVMLVWGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGGE 120
```

```
Db      61 KLGWLVGNGNSTLTGVIANREGISWATDKIQOANVFGLTOASAIRVGSFOGEE 120
Qy      121 IYAPFKSLIPMNPPDIFVGGWDISNMNLADAMAAKVPFDIDLOKOLRPYMSMLPLPGI 180
Db      121 IYAPFKSLIPMNPPDIFVGGWDISNMNLADAMAAKVPFDIDLOKOLRPYMSMLPLPGI 180
Qy      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
Db      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
Qy      241 VGLNTMTENLLAUVDRNEAISPSTLYAIACMENVPFINSPOPTFVPGILDIAIANT 300
Db      241 VGLNTMTENLLAUVDRNEAISPSTLYAIACMENVPFINSPOPTFVPGILDIAIANT 300
Qy      301 LIGGDDFKSGOTKMSVLVDPLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
Db      301 LIGGDDFKSGOTKMSVLVDPLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
Qy      361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Db      361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy      421 EDSLAAPIIIDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKPLVPFGTPPVN 480
Db      421 EDSLAAPIIIDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKPLVPFGTPPVN 480
Qy      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

```
RESULT 6
US-10-025-003-2
; Sequence 2, Application US/10025003
```

```
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-2
```

```
Query Match      100.0%; Score 510; DB 26; Length 510;
```

```
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 MFIEENFKVCEPNVKYKTEIEIOSVYNYETTELVEHNRNGTYOMIVPKSVKYEKTNIHVP 60
Db      1 MFIEENFKVCEPNVKYKTEIEIOSVYNYETTELVEHNRNGTYOMIVPKSVKYEKTNIHVP 60
Qy      61 KLGWLVGNGNSTLTGVIANREGISWATDKIQOANVFGLTOASAIRVGSFOGEE 120
Db      61 KLGWLVGNGNSTLTGVIANREGISWATDKIQOANVFGLTOASAIRVGSFOGEE 120
Qy      121 IYAPFKSLIPMNPPDIFVGGWDISNMNLADAMAAKVPFDIDLOKOLRPYMSMLPLPGI 180
Db      121 IYAPFKSLIPMNPPDIFVGGWDISNMNLADAMAAKVPFDIDLOKOLRPYMSMLPLPGI 180
Qy      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
Db      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
```

```
Db      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
Qy      241 VGLNTMTENLLAUVDRNEAISPSTLYAIACMENVPFINSPOPTFVPGILDIAIANT 300
Db      241 VGLNTMTENLLAUVDRNEAISPSTLYAIACMENVPFINSPOPTFVPGILDIAIANT 300
Qy      301 LIGGDDFKSGOTKMSVLVDPLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
Db      301 LIGGDDFKSGOTKMSVLVDPLVGAGIKPTSIIVSYNHLGNNDGMNLSAPQFRSKEISKS 360
Qy      361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Db      361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy      421 EDSLAAPIIIDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKPLVPFGTPPVN 480
Db      421 EDSLAAPIIIDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTAKPLVPFGTPPVN 480
Qy      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

```
RESULT 7
US-10-025-003-10
; Sequence 10, Application US/10025003
```

```
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-10
```

```
Query Match      100.0%; Score 510; DB 26; Length 510;
```

```
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 MFIEENFKVCEPNVKYKTEIEIOSVYNYETTELVEHNRNGTYOMIVPKSVKYEKTNIHVP 60
Db      1 MFIEENFKVCEPNVKYKTEIEIOSVYNYETTELVEHNRNGTYOMIVPKSVKYEKTNIHVP 60
Qy      61 KLGWLVGNGNSTLTGVIANREGISWATDKIQOANVFGLTOASAIRVGSFOGEE 120
Db      61 KLGWLVGNGNSTLTGVIANREGISWATDKIQOANVFGLTOASAIRVGSFOGEE 120
Qy      121 IYAPFKSLIPMNPPDIFVGGWDISNMNLADAMAAKVPFDIDLOKOLRPYMSMLPLPGI 180
Db      121 IYAPFKSLIPMNPPDIFVGGWDISNMNLADAMAAKVPFDIDLOKOLRPYMSMLPLPGI 180
Qy      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
Db      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIFKFKATKVDKVVVLTANTERYSNLV 240
Qy      241 VGLNTMTENLLAUVDRNEAISPSTLYAIACMENVPFINSPOPTFVPGILDIAIANT 300
Db      241 VGLNTMTENLLAUVDRNEAISPSTLYAIACMENVPFINSPOPTFVPGILDIAIANT 300
```

```

Qy 301 LIGDDPKSGQTKMKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNLSAPOTFRSKSIS 360
Db 301 LIGDDPKSGQTKMKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNLSAPOTFRSKSIS 360
Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIILDLVLLAEISTRIEFKAENEGKFSHPVATILSYLTAKAPLVPGTPPVN 480
Db 421 EDSLAAPIILDLVLLAEISTRIEFKAENEGKFSHPVATILSYLTAKAPLVPGTPPVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 8
US-10-718-952-2
; Sequence 2, Application US/10718952
; GENERAL INFORMATION:
; APPLICANT: Hitec, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-2

Query Match      100.0%; Score 510; DB 33; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFIEBPKVBCPNVKYKTEIEIQSVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Db 1 MFIEBPKVBCPNVKYKTEIEIQSVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Qy 61 KIGVNLVGVGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
Db 61 KIGVNLVGVGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFSLPLPMVNPDIYFGGMDISNMNLADAMARAKVFDIDLOKOLRPMESMLPLPGI 180
Db 121 IYAPFSLPLPMVNPDIYFGGMDISNMNLADAMARAKVFDIDLOKOLRPMESMLPLPGI 180
Qy 181 YPDFIAANQERANNVIKGTQEQVOQIIKDIKAFKATKYDKVNVVLTANTERSNLV 240
Db 181 YPDFIAANQERANNVIKGTQEQVOQIIKDIKAFKATKYDKVNVVLTANTERSNLV 240
Qy 241 VGLNDTMENTLLAAVDNRNEAISPSTLYAIACWENVPFINGSPONTFVGLIDLAIARNT 300
Db 241 VGLNDTMENTLLAAVDNRNEAISPSTLYAIACWENVPFINGSPONTFVGLIDLAIARNT 300
Qy 301 LIGDDPKSGQTKMKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNLSAPOTFRSKSIS 360
Db 301 LIGDDPKSGQTKMKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNLSAPOTFRSKSIS 360
Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKSTIVLHNTC 420

```

```

Qy 421 EDSLAAPIILDLVLLAEISTRIEFKAENEGKFSHPVATILSYLTAKAPLVPGTPPVN 480
Db 421 EDSLAAPIILDLVLLAEISTRIEFKAENEGKFSHPVATILSYLTAKAPLVPGTPPVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 9
US-10-718-952-10
; Sequence 10, Application US/10718952
; GENERAL INFORMATION:
; APPLICANT: Hitec, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-10

Query Match      100.0%; Score 510; DB 33; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFIEBPKVBCPNVKYKTEIEIQSVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Db 1 MFIEBPKVBCPNVKYKTEIEIQSVYNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Qy 61 KIGVNLVGVGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
Db 61 KIGVNLVGVGNGNGSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFSLPLPMVNPDIYFGGMDISNMNLADAMARAKVFDIDLOKOLRPMESMLPLPGI 180
Db 121 IYAPFSLPLPMVNPDIYFGGMDISNMNLADAMARAKVFDIDLOKOLRPMESMLPLPGI 180
Qy 181 YPDFIAANQERANNVIKGTQEQVOQIIKDIKAFKATKYDKVNVVLTANTERSNLV 240
Db 181 YPDFIAANQERANNVIKGTQEQVOQIIKDIKAFKATKYDKVNVVLTANTERSNLV 240
Qy 241 VGLNDTMENTLLAAVDNRNEAISPSTLYAIACWENVPFINGSPONTFVGLIDLAIARNT 300
Db 241 VGLNDTMENTLLAAVDNRNEAISPSTLYAIACWENVPFINGSPONTFVGLIDLAIARNT 300
Qy 301 LIGDDPKSGQTKMKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNLSAPOTFRSKSIS 360
Db 301 LIGDDPKSGQTKMKSVLVDPLVAGIKPTSIYSYNHLGNNDGMNLSAPOTFRSKSIS 360
Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIILDLVLLAEISTRIEFKAENEGKFSHPVATILSYLTAKAPLVPGTPPVN 480
Db 421 EDSLAAPIILDLVLLAEISTRIEFKAENEGKFSHPVATILSYLTAKAPLVPGTPPVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

Db 481 ALSKORAMLENTIMRACVGLAPENNMLEYK 510

RESULT 10

US-60-592-978-4180

Sequence 4180, Application US/60592978

GENERAL INFORMATION:

APPLICANT: Abad, Mark S.

TITLE OF INVENTION: Genes and Uses for Plant Improvement

FILE REFERENCE: 38-21(53452)A

CURRENT APPLICATION NUMBER: US/60/592,978

CURRENT FILING DATE: 2004-07-31

NUMBER OF SEQ ID NOS: 22568

SEQ ID NO 4180

LENGTH: 510

TYPE: PRT

ORGANISM: Glycine max

US-60-592-978-4180

Query Match 100.0%; Score 510; DB 37; Length 510;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIENFKVCEPVKYTEIETIQSVYNYETTELVEHNRNGTYQMIYKPKSVKYEFTKNIHVP 60

Db 1 MEIENFKVCEPVKYTEIETIQSVYNYETTELVEHNRNGTYQMIYKPKSVKYEFTKNIHVP 60

QY 61 KLGWMLVGMGNGNSTLTGCVIANREGISWATKDKIQQANFGSLTQASAIRVGSFQGBE 120

Db 61 KLGWMLVGMGNGNSTLTGCVIANREGISWATKDKIQQANFGSLTQASAIRVGSFQGBE 120

QY 121 IYAPFKSLPMPVNPDDIVFGGWDISNMNLADAMARAKVFDIDLQKOLRPYMSMLPLPGI 180

Db 121 IYAPFKSLPMPVNPDDIVFGGWDISNMNLADAMARAKVFDIDLQKOLRPYMSMLPLPGI 180

QY 181 YPDPFIANQOEBRANNVIKGTQEQVOQIIKDIAKFAKATKVDKVVVLTANTERYSNLY 240

Db 181 YPDPFIANQOEBRANNVIKGTQEQVOQIIKDIAKFAKATKVDKVVVLTANTERYSNLY 240

QY 241 VGLNDTMENLLAAVDNRNBAEISPTLYAIACWENVPIFGSPONTFVPGGLDILAIAANT 300

Db 241 VGLNDTMENLLAAVDNRNBAEISPTLYAIACWENVPIFGSPONTFVPGGLDILAIAANT 300

QY 301 LIGGDDPKSGQTKMSVLVDPLVAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360

Db 301 LIGGDDPKSGQTKMSVLVDPLVAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360

QY 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC 420

Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC 420

QY 421 EDSLAAPIILDLVLAELSTRIEFKANEGKFSFHPVATILSYLTAKPLVPPTPVVN 480

Db 421 EDSLAAPIILDLVLAELSTRIEFKANEGKFSFHPVATILSYLTAKPLVPPTPVVN 480

QY 481 ALSKORAMLENTIMRACVGLAPENNMLEYK 510

Db 481 ALSKORAMLENTIMRACVGLAPENNMLEYK 510

RESULT 11

US-60-592-978-6132

Sequence 6132, Application US/60592978

GENERAL INFORMATION:

APPLICANT: Abad, Mark S.

TITLE OF INVENTION: Genes and Uses for Plant Improvement

FILE REFERENCE: 38-21(53452)A

CURRENT APPLICATION NUMBER: US/60/592,978

CURRENT FILING DATE: 2004-07-31

NUMBER OF SEQ ID NOS: 22568

SEQ ID NO 6132

LENGTH: 510

TYPE: PRT

ORGANISM: Glycine max

US-60-592-978-6132

Query Match 100.0%; Score 510; DB 37; Length 510;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEIENFKVCEPVKYTEIETIQSVYNYETTELVEHNRNGTYQMIYKPKSVKYEFTKNIHVP 60

Db 1 MEIENFKVCEPVKYTEIETIQSVYNYETTELVEHNRNGTYQMIYKPKSVKYEFTKNIHVP 60

QY 61 KLGWMLVGMGNGNSTLTGCVIANREGISWATKDKIQQANFGSLTQASAIRVGSFQGBE 120

Db 61 KLGWMLVGMGNGNSTLTGCVIANREGISWATKDKIQQANFGSLTQASAIRVGSFQGBE 120

QY 121 IYAPFKSLPMPVNPDDIVFGGWDISNMNLADAMARAKVFDIDLQKOLRPYMSMLPLPGI 180

Db 121 IYAPFKSLPMPVNPDDIVFGGWDISNMNLADAMARAKVFDIDLQKOLRPYMSMLPLPGI 180

QY 181 YPDPFIANQOEBRANNVIKGTQEQVOQIIKDIAKFAKATKVDKVVVLTANTERYSNLY 240

Db 181 YPDPFIANQOEBRANNVIKGTQEQVOQIIKDIAKFAKATKVDKVVVLTANTERYSNLY 240

QY 241 VGLNDTMENLLAAVDNRNBAEISPTLYAIACWENVPIFGSPONTFVPGGLDILAIAANT 300

Db 241 VGLNDTMENLLAAVDNRNBAEISPTLYAIACWENVPIFGSPONTFVPGGLDILAIAANT 300

QY 301 LIGGDDPKSGQTKMSVLVDPLVAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360

Db 301 LIGGDDPKSGQTKMSVLVDPLVAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360

QY 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC 420

Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMDEYTSIEIFMGKSTIVLHNTC 420

QY 421 EDSLAAPIILDLVLAELSTRIEFKANEGKFSFHPVATILSYLTAKPLVPPTPVVN 480

Db 421 EDSLAAPIILDLVLAELSTRIEFKANEGKFSFHPVATILSYLTAKPLVPPTPVVN 480

QY 481 ALSKORAMLENTIMRACVGLAPENNMLEYK 510

Db 481 ALSKORAMLENTIMRACVGLAPENNMLEYK 510

RESULT 12

US-10-424-599-213009

Sequence 213009, Application US/10424599

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 213009

LENGTH: 511

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(511)

OTHER INFORMATION: unsure at all Xaa locations

OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1.dep

US-10-424-599-213009

Query Match 86.3%; Score 440; DB 30; Length 511;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 GNGSLTGTGVIANREGISWATKIOQANYFSLAQASAIRVSGFGEETVAFKSLP 130
 DB 72 GNGSLTGTGVIANREGISWATKIOQANYFSLAQASAIRVSGFGEETVAFKSLP 131
 QY 131 MNPDDIVFGWMDISNNMLADAMARAKVPDIDLOKOLRPMESMLPLPGIYDDPFIANQ 190
 DB 132 MNPDDIVFGWMDISNNMLADAMARAKVPDIDLOKOLRPMESMLPLPGIYDDPFIANQ 191
 QY 191 EEPANNVIKTKQBOVOQIIKDIAKFAKATKVDKVVLTANTERSNLVGLNDTMENT 250
 DB 192 EEPANNVIKTKQBOVOQIIKDIAKFAKATKVDKVVLTANTERSNLVGLNDTMENT 251
 QY 251 LAADVNRNEAISPSTLYAIAACWENVPFINGSPONTFVPGILDIAIARNTLIGDDPKSG 310
 DB 252 LAADVNRNEAISPSTLYAIAACWENVPFINGSPONTFVPGILDIAIARNTLIGDDPKSG 311
 QY 311 QTKMKSVDPLVAGAGIKPTSIYSYNHLGNNDGMLSAPOTFRSKEISKNVDDMVNSN 370
 DB 312 QTKMKSVDPLVAGAGIKPTSIYSYNHLGNNDGMLSAPOTFRSKEISKNVDDMVNSN 371
 QY 371 AIIYEGEHPDHVVVITKVPYVDSKRAMDEYTSSEIFMGKSTI VLHNTCEDSLAAPIT 430
 DB 372 AIIYEGEHPDHVVVITKVPYVDSKRAMDEYTSSEIFMGKSTI VLHNTCEDSLAAPIT 431
 QY 431 LDVLVLAELSTRIEFKAENEGKFSHPVATITSLTKAPLVPGTPVNNALSKORAMLE 490
 DB 432 LDVLVLAELSTRIEFKAENEGKFSHPVATITSLTKAPLVPGTPVNNALSKORAMLE 491
 QY 491 NIMRACVGLAPENNMLEYK 510
 DB 492 NIMRACVGLAPENNMLEYK 511

RESULT 13

US-08-835-751A-6
 ; Sequence 6, Application US/08835751A

GENERAL INFORMATION:

APPLICANT: HITZ, WILLIAM D.

TITLE OF INVENTION: SOYBEAN PLANTS PRODUCING SEEDS WITH

TITLE OF INVENTION: REDUCED LEVELS OF RAFFINOSE

TITLE OF INVENTION: SACCARIDES AND PHYTIC ACID

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

STREET: 1007 MARKET STREET

CITY: WILMINGTON

STATE: DELAWARE

COUNTRY: UNITED STATES OF AMERICA

ZIP: 19898

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.50 INCH

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: MICROSOFT WINDOWS 95

SOFTWARE: MICROSOFT WORD FOR WINDOWS 95 (7.0)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/835, 751A

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: MAJARIAN, WILLIAM R.

REGISTRATION NUMBER: P41,173

REFERENCE/DOCKET NUMBER: BB-1077

TELEPHONE: (302)992-4926

TELEFAX: (302)773-0164

SEQUENCE CHARACTERISTICS:

LENGTH: 510 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-835-751A-6

Query Match 80.2%; Score 409; DB 12; Length 510;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFLENKVECPNKKYETETIQSVYNYETTELHNNNGTYQWIVKPSVYKEEKTNIHP 60
 DB 1 MFLENKVECPNKKYETETIQSVYNYETTELHNNNGTYQWIVKPSVYKEEKTNIHP 60
 QY 61 KLGVMVLVGMGNGSLTGTGVIANREGISWATKIOQANYFSLAQASAIRVSGFGEET 120
 DB 61 KLGVMVLVGMGNGSLTGTGVIANREGISWATKIOQANYFSLAQASAIRVSGFGEET 120
 QY 121 IYAPFKSLPMPVNPDDIVFGWMDISNNMLADAMARAKVPDIDLOKOLRPMESMLPLPGI 180
 DB 121 IYAPFKSLPMPVNPDDIVFGWMDISNNMLADAMARAKVPDIDLOKOLRPMESMLPLPGI 180
 QY 181 YDPDFIAANQEEERANNVVKGTQBOVOQIIKDIAKFAKATKVDKVVLTANTERSNLV 240
 DB 181 YDPDFIAANQEEERANNVVKGTQBOVOQIIKDIAKFAKATKVDKVVLTANTERSNLV 240
 QY 241 VGLNDTMENTLAADVNRNEAISPSTLYAIAACWENVPFINGSPONTFVPGILDIAIARNT 300
 DB 241 VGLNDTMENTLAADVNRNEAISPSTLYAIAACWENVPFINGSPONTFVPGILDIAIARNT 300
 QY 301 LIGDDPKSGQTKMKSVDPLVAGAGIKPTSIYSYNHLGNNDGMLSAPOTFRSKEISKS 360
 DB 301 LIGDDPKSGQTKMKSVDPLVAGAGIKPTSIYSYNHLGNNDGMLSAPOTFRSKEISKS 360
 QY 361 NVYDDMVNSNALIYEGEHPDHVVVITKVPYVDSKRAMDEYTSSEIFMGKSTI VLHNTC 420
 DB 361 NVYDDMVNSNALIYEGEHPDHVVVITKVPYVDSKRAMDEYTSSEIFMGKSTI VLHNTC 420
 QY 421 EDSLAPITIIDVLVLAELSTRIEFKAENEGKFSHPVATITSLTKAPLVPGTPVNNAL 480
 DB 421 EDSLAPITIIDVLVLAELSTRIEFKAENEGKFSHPVATITSLTKAPLVPGTPVNNAL 480
 QY 481 ALSKORAMLENIWRACVGLAPENNMLEYK 510
 DB 481 ALSKORAMLENIWRACVGLAPENNMLEYK 510

RESULT 14

US-09-299-315-6

; Sequence 6, Application US/09299315

GENERAL INFORMATION:

APPLICANT: Hitz, William

APPLICANT: Sebastian, Scott

APPLICANT: Grace, John

TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE

TITLE OF INVENTION: SACCARIDES AND PHYTIC ACID

FILE REFERENCE: BB-1077-B

CURRENT APPLICATION NUMBER: US/09/299,315

CURRENT FILING DATE: 1999-04-26

EARLIER APPLICATION NUMBER: 08/835,751

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Microsoft Office 97

SEQ ID NO 6

LENGTH: 510

TYPE: PRT

ORGANISM: Glycine max

US-09-299-315-6

Query Match 80.2%; Score 409; DB 16; Length 510;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 MFLENKVECPNKKYETETIQSVYNYETTELHNNNGTYQWIVKPSVYKEEKTNIHP 60

```

Db      1 MEIENFKVCEPNVKTETEIQSVYNYETTELHNNNGTYOWI VPKSVKYEKTNHVP 60
Qy      61 KLGVMVVGKNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIYVGSFQGBE 120
Db      61 KLGVMVVGKNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIYVGSFQGBE 120
Qy      121 IYAPFKSLPVMNPPDIVFGWDISNMNLADAMARAKVFDIDLQKQLRPYMESMLPLPGI 180
Db      121 IYAPFKSLPVMNPPDIVFGWDISNMNLADAMARAKVFDIDLQKQLRPYMESMLPLPGI 180
Qy      181 YPDPFIANNOEBRANNVIGTKQEOVOQI IKDIAKFEATKVDKVVLTANTERSNLY 240
Db      181 YPDPFIANNOEBRANNVIGTKQEOVOQI IKDIAKFEATKVDKVVLTANTERSNLY 240
Qy      241 VGLNDTMENLLAAVDRENEAISPSTLYAIACWENVPIFGSPONTFVPGILDLAIARNT 300
Db      241 VGLNDTMENLLAAVDRENEAISPSTLYAIACWENVPIFGSPONTFVPGILDLAIARNT 300
Qy      301 LIGDDFFKSGQTKMSVLDVFLVAGIKPTISVSYNHLGNDGMNLSAPQTFRSKEISKS 360
Db      301 LIGDDFFKSGQTKMSVLDVFLVAGIKPTISVSYNHLGNDGMNLSAPQTFRSKEISKS 360
Qy      361 NVDDMVNSNALLYEPGEHPDHVVYIKVVPYVDSKRAMDEYTSIIFMGKSTIYLAHNTC 420
Db      361 NVDDMVNSNALLYEPGEHPDHVVYIKVVPYVDSKRAMDEYTSIIFMGKSTIYLAHNTC 420
Qy      421 EDSLAAPIIIDLVLAEISTRIEFKAENEGKFSFHPVATILSYLTAKPLVPPTPVVN 480
Db      421 EDSLAAPIIIDLVLAEISTRIEFKAENEGKFSFHPVATILSYLTAKPLVPPTPVVN 480
Qy      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

```

RESULT 15
US-09-304-534-6
; Sequence 6, Application US/09304534
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebaetian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/09/304,534
; EARLIER FILING DATE: 1999-05-04
; EARLIER APPLICATION NUMBER: 08/835,751
; EARLIER FILING DATE: APRIL 8, 1997
; EARLIER APPLICATION NUMBER: PCT/US98/06822
; EARLIER FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-09-304-534-6

```

```

Query Match      80.2%; Score 409; DB 17; Length 510;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 MEIENFKVCEPNVKTETEIQSVYNYETTELHNNNGTYOWI VPKSVKYEKTNHVP 60
Db      1 MEIENFKVCEPNVKTETEIQSVYNYETTELHNNNGTYOWI VPKSVKYEKTNHVP 60
Qy      61 KLGVMVVGKNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIYVGSFQGBE 120
Db      61 KLGVMVVGKNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIYVGSFQGBE 120
Qy      121 IYAPFKSLPVMNPPDIVFGWDISNMNLADAMARAKVFDIDLQKQLRPYMESMLPLPGI 180

```

```

Db      121 IYAPFKSLPVMNPPDIVFGWDISNMNLADAMARAKVFDIDLQKQLRPYMESMLPLPGI 180
Qy      181 YPDPFIANNOEBRANNVIGTKQEOVOQI IKDIAKFEATKVDKVVLTANTERSNLY 240
Db      181 YPDPFIANNOEBRANNVIGTKQEOVOQI IKDIAKFEATKVDKVVLTANTERSNLY 240
Qy      241 VGLNDTMENLLAAVDRENEAISPSTLYAIACWENVPIFGSPONTFVPGILDLAIARNT 300
Db      241 VGLNDTMENLLAAVDRENEAISPSTLYAIACWENVPIFGSPONTFVPGILDLAIARNT 300
Qy      301 LIGDDFFKSGQTKMSVLDVFLVAGIKPTISVSYNHLGNDGMNLSAPQTFRSKEISKS 360
Db      301 LIGDDFFKSGQTKMSVLDVFLVAGIKPTISVSYNHLGNDGMNLSAPQTFRSKEISKS 360
Qy      361 NVDDMVNSNALLYEPGEHPDHVVYIKVVPYVDSKRAMDEYTSIIFMGKSTIYLAHNTC 420
Db      361 NVDDMVNSNALLYEPGEHPDHVVYIKVVPYVDSKRAMDEYTSIIFMGKSTIYLAHNTC 420
Qy      421 EDSLAAPIIIDLVLAEISTRIEFKAENEGKFSFHPVATILSYLTAKPLVPPTPVVN 480
Db      421 EDSLAAPIIIDLVLAEISTRIEFKAENEGKFSFHPVATILSYLTAKPLVPPTPVVN 480
Qy      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

```

Search completed: June 7, 2005, 17:21:12
Job time : 278.5 secs

```



```
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Masucci, James D.
/ APPLICANT: Zhou, Yihua
/ TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53374)B
/ CURRENT APPLICATION NUMBER: US/10/703,032
/ CURRENT FILING DATE: 2003-11-06
/ PRIOR APPLICATION NUMBER: 10/020,338
/ PRIOR FILING DATE: 2001-12-12
/ NUMBER OF SEQ ID NOS: 21164
/ SEQ ID NO 108933
/ LENGTH: 192
/ TYPE: PRT
/ ORGANISM: Triticum aestivum
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_TA_3351.pep
US-10-703-032-108933
```

```
Query Match      13.1%; Score 67; DB 6; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.2e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNDGMNLAPQTFRSKEISKS 360
DB      99 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNDGMNLAPQTFRSKEISKS 158
QY      361 NVVDDMV 367
DB      159 NVVDDMV 165
```

RESULT 3

```
US-10-703-032-135209
/ Sequence 135209, Application US/10703032
/ GENERAL INFORMATION:
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Andersen, Scott E.
/ APPLICANT: Byrum, Joseph R.
/ APPLICANT: Comer, Timothy W.
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Masucci, James D.
/ APPLICANT: Zhou, Yihua
/ TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53374)B
/ CURRENT APPLICATION NUMBER: US/10/703,032
/ CURRENT FILING DATE: 2003-11-06
/ PRIOR APPLICATION NUMBER: 10/020,338
/ PRIOR FILING DATE: 2001-12-12
/ NUMBER OF SEQ ID NOS: 21164
/ SEQ ID NO 135209
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Triticum aestivum
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_TA_29627.pep
US-10-703-032-135209
```

```
Query Match      13.1%; Score 67; DB 6; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.2e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNDGMNLAPQTFRSKEISKS 360
DB      301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNDGMNLAPQTFRSKEISKS 360
QY      361 NVVDDMV 367
DB      361 NVVDDMV 367
```

RESULT 4

```
US-11-064-295-4
/ Sequence 4, Application US/11064295
/ GENERAL INFORMATION:
/ APPLICANT: Liebergessel, Matthias
/ APPLICANT: Shi, Jintui
/ APPLICANT: Singletary, George
/ TITLE OF INVENTION: MODULATING MYO-INOSITOL CATABOLISM IN
/ TITLE OF INVENTION: PLANTS
/ FILE REFERENCE: 35718/287607
/ CURRENT APPLICATION NUMBER: US/11/064,295
/ CURRENT FILING DATE: 2005-02-23
/ PRIOR APPLICATION NUMBER: 60/547,640
/ PRIOR FILING DATE: 2004-02-25
/ NUMBER OF SEQ ID NOS: 9
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Zea mays
US-11-064-295-4
```

```
Query Match      13.1%; Score 67; DB 7; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.2e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNDGMNLAPQTFRSKEISKS 360
DB      301 LIGDDFKSGQTKMSVLVDPLVGAGIKPTSIIVSYNHLGNDGMNLAPQTFRSKEISKS 360
QY      361 NVVDDMV 367
DB      361 NVVDDMV 367
```

RESULT 5

```
US-11-097-143-23283
/ Sequence 23283, Application US/11097143
/ GENERAL INFORMATION:
/ APPLICANT: Venetier, J. Craig
/ APPLICANT: et al.
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ FILE REFERENCE: CL000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ CURRENT FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 23283
/ LENGTH: 565
/ TYPE: PRT
/ ORGANISM: DROSOPHILA
US-11-097-143-23283
```

```
Query Match      5.7%; Score 29; DB 7; Length 565;
Best Local Similarity 100.0%; Pred. No. 5.7e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	378	EHPDHVVIKYVPYVGDSKRA	DEYTS	406
Db	379	EHPDHVVIKYVPYVGDSKRA	DEYTS	407

RESULT 6

```

US-10-990-328A-8846
: Sequence 8846, Application US/10990328A
:
: GENERAL INFORMATION:
:
: APPLICANT: CARCILLU, Michele
:
: TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
:
: TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
:
: TITLE OF INVENTION: USES THEREOF
:
: FILE REFERENCE: CL001495
:
: CURRENT APPLICATION NUMBER: US/10/990,328A
:
: CURRENT FILING DATE: 2004-11-17
:
: NUMBER OF SEQ ID NOS: 558824
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 8846
:
: LENGTH: 235
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
:
: US-10-990-328A-8846

```

Query Match 3.7%; Score 19; DB 6; Length 235;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      415 VLHNTCEDSLAAPIIDL 433
          |||||
Db      50 VLHNTCEDSLAAPIIDL 68
```

RESULT 7

```

US-10-703-032-147024
? Sequence 147024, Application US/10703032
? GENERAL INFORMATION:
? APPLICANT: Kovalic, David K.
? APPLICANT: Andersen, Scott E.
? APPLICANT: Byrum, Joseph R.
? APPLICANT: Conner, Timothy W.
? APPLICANT: Cao, Yongwei
? APPLICANT: Maesucci, James D.
? APPLICANT: Zhou, Yihua
? TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
? FILE OF INVENTION: Plants
? FILE REFERENCE: 38-21(53374)B
? CURRENT APPLICATION NUMBER: US/10/703,032
? CURRENT FILING DATE: 2003-11-06
? PRIOR APPLICATION NUMBER: 10/020,338
? PRIOR FILING DATE: 2001-12-12
? NUMBER OF SEQ ID NOS: 211164
? SEQ ID NO 147024
? LENGTH: 132
? TYPE: PR1
? ORGANISM: Triticum aestivum
? FEATURE:
? OTHER INFORMATION: Clone ID: PAT_TA_41442.pep
? US-10-703-032-147024

```

Query Match 3.5%; Score 18; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      221 KVDKVVWLMTANTERYSN 238
          |||||
Db      105 KVDKVVWLMTANTERYSN 122
```

RESULT 8
US-10-450-763-34728
; Sequence 34728, Application US/10450763

```

: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc
: TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
: FILE REFERENCE: 790C1P3/US
: CURRENT APPLICATION NUMBER: US/10/450,763
: PRIOR FILING DATE: 2003-06-11
: PRIOR APPLICATION NUMBER: PCT/US01/08631
: PRIOR FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: 09/540,217
: PRIOR FILING DATE: 2000-03-31
: PRIOR APPLICATION NUMBER: 09/649,167
: PRIOR FILING DATE: 2000-08-23
: NUMBER OF SEQ ID NOS: 60736
: SOFTWARE: Custom
: SEQ ID NO 34728
: LENGTH: 111
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-450-763-34728

```

Query Match Similarity	3.3%	Score 17	DB 6	length 111
Best Local Similarity	100.0%	Pred. No.	2.9e-03	
Matches 17	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	62	LGVM LVGM GNGN NSTLT	78
Db	76	LGVM LVGM GNGN NSTLT	92

RESULT 9

US-60-669-175-37019
Sequence 37019, Application US/60669175
GENERAL INFORMATION:
APPLICANT: Monsanto Technology LLC
APPLICANT: Baum, James A
APPLICANT: Gilbertson, Larry A
APPLICANT: Kovalic, David K
APPLICANT: Larosa, Thomas J
APPLICANT: Lu, Maolong
APPLICANT: Munyikwa, Tichifa R. I.
APPLICANT: Roberts, James K
APPLICANT: Wu, Wei
APPLICANT: Zhang, Bei
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CONTROL OF INSECT INFESTATION IN PLANT
FILE REFERENCE: 38-21(53597)
CURRENT APPLICATION NUMBER: US/60/669,175
CURRENT FILING DATE: 2005-04-07
PRIOR APPLICATION NUMBER: 60560842
PRIOR FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: 60565632
PRIOR FILING DATE: 2004-04-27
PRIOR APPLICATION NUMBER: 60579062
PRIOR FILING DATE: 2004-06-11
PRIOR APPLICATION NUMBER: 60603421
PRIOR FILING DATE: 2004-08-20
PRIOR APPLICATION NUMBER: 60617261
PRIOR FILING DATE: 2004-10-11
NUMBER OF SEQ ID NOS: 40774
SEQ ID NO 37019

```

? LENGTH: 280
? TYPE: prt
? ORGANISM: Diabrotica virgifera
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (239)..(239)
? OTHER INFORMATION: Xaa can be any naturally occurring amino acid
? FEATURE:
? OTHER INFORMATION: Coding regions on vCDNA: vCDNA=SeqID_16716; Strand=+; Position=1-{
? FEATURE:
? OTHER INFORMATION: Homolog annotation: Hit_ID=XP_214319.2; Match level="QueryCoverage
? OTHER INFORMATION: =99%, HitCoverage=50%, E-value=9e-82, Identity=55%"; Hit description:
? OTHER INFORMATION: =similar to myo-inositol 1-phosphate synthase A1 [Rattus norvegicus]
? FEATURE:

```

OTHER INFORMATION: Gene Ontology Annotation: Molecular function=inositol-3-phosphate
OTHER INFORMATION: synthase activity; Cellular component=cellular component unknown;
OTHER INFORMATION: Biological process=Myo-inositol metabolism
FEATURE:
OTHER INFORMATION: Pfam annotation: Pfam_ID=Inos-1-P_synth, Match level=Score=33.5,
OTHER INFORMATION: -value=5.8e-12, Copies=1"; Pfam description=Myo-inositol-1-phosph
US-60-669-175-37019

Query Match 3.3%; Score 17; DB 8; Length 280;
Best Local Similarity 100.0%; Pred. No. 7.1e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 LGVMLVGMGNGNGST 76
DB 50 PRLGVMVLVGMGNGNGST 66

RESULT 10
US-10-990-328A-10375
; Sequence 10375, Application US/10990328A

GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
FILE REFERENCE: CL001495
CURRENT APPLICATION NUMBER: US/10/990,328A
CURRENT FILING DATE: 2004-11-17
NUMBER OF SEQ ID NOS: 558824
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10375
LENGTH: 558
TYPE: PRT
ORGANISM: Homo sapiens
US-10-990-328A-10375

Query Match 3.3%; Score 17; DB 6; Length 558;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 LGVMLVGMGNGNGSTLT 78
DB 59 LGVMLVGMGNGNGSTLT 75

RESULT 11
US-60-671-501-362
; Sequence 362, Application US/60671501
GENERAL INFORMATION:
APPLICANT: Wm Van Crielinge
APPLICANT: Josef Straub
TITLE OF INVENTION: METHYLATION MARKERS FOR DIAGNOSIS AND
TITLE OF INVENTION: TREATMENT OF CANCERS
FILE REFERENCE: 006791.00001
CURRENT APPLICATION NUMBER: US/60/671,501
CURRENT FILING DATE: 2005-04-15
NUMBER OF SEQ ID NOS: 420
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 362
LENGTH: 559
TYPE: PRT
ORGANISM: Homo sapiens
US-60-671-501-362

Query Match 3.3%; Score 17; DB 8; Length 559;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 LGVMLVGMGNGNGSTLT 78
DB 60 LGVMLVGMGNGNGSTLT 76

RESULT 12
US-10-450-763-52092
; Sequence 52092, Application US/10450763

GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 52092
LENGTH: 704
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (224)..(644)
OTHER INFORMATION: Myo-inositol-1-phosphate synthase domain identified by Pfam,
OTHER INFORMATION: accession name Inos-1-P_synth, E-value=1.9e-68, Pfam score of 240.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(704)
OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-52092

Query Match 2.9%; Score 15; DB 6; Length 704;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 VLAHNTCEDSLAAPI 429
DB 549 VLAHNTCEDSLAAPI 563

RESULT 13
US-10-450-763-34498
; Sequence 34498, Application US/10450763
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 34498
LENGTH: 237
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (1)..(237)
OTHER INFORMATION: Myo-inositol-1-phosphate synthase domain identified by Pfam,
OTHER INFORMATION: accession name Inos-1-P_synth, E-value=3.9e-07, Pfam score of -73.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(237)
OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-34498

Query Match 2.5%; Score 13; DB 6; Length 237;
Best Local Similarity 100.0%; Pred. No. 8.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 331 SIVSYNHLGNNDG 343
DB 170 SIVSYNHLGNNDG 182

RESULT 14
US-10-450-763-34497
; Sequence 34497, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/549,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 34497
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(244)
; OTHER INFORMATION: Myo-inositol-1-phosphate synthase domain identified by Pfam,
US-10-450-763-34497

Query Match 2.5%; Score 13; DB 6; Length 244;
Best Local Similarity 100.0%; Pred. No. 8.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 SIVSYNHLGNNDG 343
DB 230 SIVSYNHLGNNDG 242

RESULT 15
US-10-450-763-52091
; Sequence 52091, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/549,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 52091
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (3)..(254)
; OTHER INFORMATION: Myo-inositol-1-phosphate synthase domain identified by Pfam,
OTHER INFORMATION: accession name Inos-1-P_synth, E-value=4.5e-27, Pfam score of 103

US-10-450-763-52091
Query Match 2.5%; Score 13; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 8.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 331 SIVSYNHLGNNDG 343
DB 240 SIVSYNHLGNNDG 252

Search completed: June 7, 2005, 17:22:47
Job time : 43.5 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:49:11 ; Search time 6689.71 Seconds
(without alignments)
11103.899 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533

Sequence: 1 atgtcaccgagaattttaa.....tgattctcgagtaacagta 1533

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hlg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_scs.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	146	9.5	1533	6	BD075269 Soybean p
2	146	9.5	1739	6	AY038802 Glycine m
3	146	9.5	1782	6	BD075266 Soybean p
4	115	7.5	1791	8	AF293970 Glycine m
5	69	4.5	2994	8	AF282263 Phaseolus
6	52	3.4	143012	8	AC125389 Medicago
7	50	3.3	62713	8	AP006418 Lotus cor
8	45	2.9	1391	8	AY005128 Actinidia
9	41	2.7	119858	8	AP006419 Lotus cor
10	32	2.1	1372	8	AB090886 Aster trl
11	32	2.1	1978	8	CPINOLG U32632 C.paradis
12	29	1.9	2053	8	MCU32511 Mesembryant
13	26	1.7	1533	8	AY323824 Xerophyta
14	26	1.7	1564	8	AY096554 Arabidops
15	26	1.7	1781	6	AX054630 Sequence
16	26	1.7	1781	8	BN066307 Brascica na
17	26	1.7	1845	8	AY028259 Avicennia
18	26	1.7	1863	8	AY065415 Arabidops
19	26	1.7	1931	8	AB032073 Nicotiana

20	26	1.7	1950	6	E27176	E27176 Novel INPS
21	26	1.7	1959	8	AB009881	AB009881 Nicotiana
22	26	1.7	1986	8	AR433879	AR433879 Suaeda ma
23	26	1.7	80088	8	AT131816	AT131816 Arabidops
24	23	1.5	35	6	BD075267	BD075267 Soybean p
25	23	1.5	1535	6	BD073472	BD073472 Regulated
26	23	1.5	1538	8	AF120146	AF120146 Triticum
27	23	1.5	1538	8	AF120147	AF120147 Triticum
28	23	1.5	1538	8	AF120148	AF120148 Triticum
29	23	1.5	1539	8	AF412340	AF412340 Porteresi
30	23	1.5	1578	6	E13612	E13612 gDNA encodi
31	23	1.5	1578	6	E17392	E17392 gDNA encodi
32	23	1.5	1665	6	BD073470	BD073470 Regulated
33	23	1.5	1665	8	AF056326	AF056326 Zea mays
34	23	1.5	1868	8	AB012107	AB012107 Oryza sat
35	23	1.5	1914	8	AK103501	AK103501 Oryza sat
36	23	1.5	1915	8	AK058750	AK058750 Oryza sat
37	23	1.5	1931	6	AR137808	AR137808 Sequence
38	23	1.5	1931	6	AR170167	AR170167 Sequence
39	23	1.5	3546	6	AR137811	AR137811 Sequence
40	23	1.5	3546	6	AR137812	AR137812 Sequence
41	23	1.5	3546	6	AR170170	AR170170 Sequence
42	23	1.5	3546	6	AR170171	AR170171 Sequence
43	23	1.5	3618	8	AF323175	AF323175 Zea mays
44	23	1.5	129658	8	AC116604	AC116604 Genomic s
45	22	1.4	682	8	AT1529402	AT1529402 Arabidops

ALIGNMENTS

RESULT 1	BD075269	1533 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	Soybean plant producing seeds with reduced levels of raffinose				
LOCUS	BD075269				
ACCESSION	BD075269.1	GI:22620872			
VERSION	JP 2001519665-A/4.				
KEYWORDS	unidentified				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1533)				
AUTHORS	Hitz,W.D. and Sebastian,S.A.				
TITLE	Soybean plant producing seeds with reduced levels of raffinose				
JOURNAL	saccharides and phytic acid				
COMMENT	Patent: JP 2001519665-A 4 23-OCT-2001; EI DU PONT DE NEMOURS AND CO OS Soybean line LR33 PN JP 2001519665-A/4 PD 23-OCT-2001 PF 07-APR-1998 JP 1998543012 PR 08-APR-1997 US 08/835751 PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN PC C12N15/52,C12N15/82,C12N15/11,C12N9/90,A01H5/00 CC Strandedness: Double; CC Topology: Linear; CC Soybean plant producing seeds with reduced levels of raffinose CC and phytic acid FH Key Location/Qualifiers FT CDS Location/Qualifiers 1..1533 1..1533. /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644"				
ORIGIN	Query Match	9.5%	Score 146;	DB 6;	Length 1533;
	Best Local Similarity	100.0%	Pred. No. 6,7e-72;		
	Matches 146;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	ATGAAAAATGTCCTTCATTAATGAAGACCCCTAGAACACTTTTGTATCCAGGGCTGATT	876
Db	ATGAAAAATGTCCTTCATTAATGAAGACCCCTAGAACACTTTTGTATCCAGGGCTGATT	876
QY	GATCTTGCATTCGAGAGAACACTTTGATTTGGTGAAGATGACTTCAAGATGTGTACAGCC	936
Db	GATCTTGCATTCGAGAGAACACTTTGATTTGGTGAAGATGACTTCAAGATGTGTACAGCC	936
QY	AAATGAAATCTGTGTTGTTGATT	962
Db	AAATGAAATCTGTGTTGTTGATT	962
RESULT 2		
AY038802	1739 bp	linear
LOCUS	Glycine max myo-inositol-1-phosphate synthase mRNA, complete cds.	
DEFINITION	AY038802	
ACCESSION	AY038802	
VERSION	GI:14764465	
KEYWORDS	Glycine max (soybean)	
SOURCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.	
ORGANISM	1 (bases 1 to 1739)	
REFERENCE	Hiltz,W.D., Carlson,T.J., Kerr,P. and Sebastian,S.	
AUTHORS	Biochemical and Molecular Characterization of a Mutation that Confers a Decreased Raffinose and Phytic Acid Phenotype on Soybean Seeds	
TITLE	Unpublished	
JOURNAL	2 (bases 1 to 1739)	
REFERENCE	Carlson,T.J. and Hiltz,W.D.	
AUTHORS	Submitted (06-JUN-2001)	
JOURNAL	Wilmington, DE 19880-0402, USA	
FEATURES	Location/Qualifiers	
SOURCE	1..1739	
	/organism="Glycine max"	
	/mol_type="mRNA"	
	/cultivar="Wye"	
	/db_xref="taxon:3847"	
	/dev_stage="developing embryo"	
	11..1543	
	/EC_number="5.5.1.4"	
	/function="cyclizes glucose 6-phosphate to L-myo-inositol-1-phosphate"	
	/note="m1PS-1A"	
	/codon_start=1	
	/product="myo-inositol-1-phosphate synthase"	
	/protein_id="AAK72098.1"	
	/db_xref="GI:14764466"	
	/translation="MFIENFVRECPNRYTETELQSVYNYVETELAHENRGTYQWTLV KPKSVKBFKTNIVPKLVGWLVGNGNSTLTGGVIANREGISMAKDKIQQANTV GSLVQASAIRGSGFQGEIYAPFKSLPMVNPDDIVFGWQISNMNLADNAARAVPDE IDLQKRLPYMESMLPLGIYDPFIANDERANNAVYKGTQKQVOQIIDIAFKFKM ATKDKVVLWLTANTERSNLVGLINDMENTLAAVBNEMEISPTSLYACVMENYV PFNGSPONTFVPGILDAIARNTLIGDDPKSGOTKMSVLDVPIVAGAGKPTISV YNHIGNDGNMLSAQPTFRKETSISNVYDVMNSNALLYRGGHPHYVIVIKVPPVYV GDSGRANDEYTSBETFMGSKSTIVLHNTDESLAAPILLDLVLAEISTRTERFAENB GKFSFHPVATILSYLTAAPIVPGTVPVNALSKORALENIMRACVGLAENNMILE YK"	
CDS		

Query Match	9.5%	Score 146	DB 8	Length 1739
Best Local Similarity	100.0%	Pred. No. 6.7e-72		
Matches 146	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	817	ATGGAATAATGTCCTTCATTATATGAGAGCCCTCAGAACACTTTTGTAACAGGGCTGATT		876
Db	827	ATGGAATAATGTCCTTCATTATATGAGAGCCCTCAGAACACTTTTGTAACAGGGCTGATT		886

QY	877	GATCTGGCATTGCGAGGAAACACTTGTATTTGGTGGAGATGACTTCAAGAGTGTGAGACC	936
Db	887	GATCTTGCATTGCGAGGAAACACTTGTATTTGGTGGAGATGACTTCAAGAGTGTGAGACC	946
QY	937	AAATGAAATCTGCTTGGTGGATTT	962
Db	947	AAATGAAATCTGCTTGGTGGATTT	972

RESULT 3
BD075266

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

DEFINITION Soybean plant producing seeds with reduced levels of raffinose saccharides and phytic acid.

ACCESSION	BD075266
VERSION	BD075266.1
KEYWORDS	GI:22620869
SOURCE	JP 2001519665-A/1.
	unidentified

ORGANISM	unidentified
REFERENCE	1 (bases 1 t

0 1782)

AUTHORS Hitz, W.D. and Sebastian, S.A.
TITLE Soybean plant producing seeds with reduced levels of raffinose saccharides and phytic acid

JOURNAL Patent : JP 2001519665-A I 23-OCT-2001,
EI DU PONT DE NEMOURS AND CO
OS Soybean line LR13

PN JP 2001519665-A/1
PD 23-OCT-2001
PF 07-APR-1998 JP 19

98543012

PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN
PC C12N15/52, C12N15/82, C12N15/11, C12N9/90, A01H5/00 CC

Strandedness: Double;
CC Topology: Linear;
CC Soybean plant producing seeds with reduced

levels of raffinose
CC saccharides
CC and phytic acid
CC

PH	key	Location/Qualifiers
FT	CDS	54. .1586.

```
FEATURES location/Qualifiers
source 1. .1782
```

```
/organism="unidentified"  
/mol_type="genomic DNA"
```

```
/db_xref="taxon:32644"
```

ORIGIN

Query Match	9.5%;	Score 146;	DB 6;	Length 1782;
Best Local Similarity	100.0%;	Pred. No. 6.7e-72;		

Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

817 ATGGAATGTTCTTTCATTATGGAAGCCCTCAGAACCTTTGTACCAGGCTGATT 876

Db 870 ATGGAATGTTCTTCATTAAAGGAGCCCTAGAACACTTTTGTACCAGGCTGATT 929

877 GATCTTCCATCGCGAGAACACTTTGATTGGTGGAGATGACTTCAAGAGTGGTCAGACC 936

930 GATCTTTCGCCATCTCCGAGGAGACACCTTTTGATTTGCTGCAAGATGACCTTTCACAGAGATGCTTCAGACC 989

[illegible]

93 / AAAAGAAACGCGTGGTGGATTT 962

Db 990 AAATGAATCTGTGTTGGTGA^{TTT} 1015

RESULT 4

AF293970

LOCUS	AF293970	1791 bp	mRNA	linear	PLN 03-MAY-200
DEFINITION	Glycine max myo-inositol-3-phosphate synthase (MIPS) mRNA, complete				

cds.
AF293970
ACCESSION

VERSION AF293970.1 GI:13936690

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS
ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 5,4e-54;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29 GTCTTAATGTAAGTACACCGAGACTGAGATTCAGTCCCTGTACAACTAGCAACACCGC 88
DB 90 GTCTTAATGTAAGTACACCGAGACTGAGATTCAGTCCCTGTACAACTAGCAACACCGC 149
QY 89 AACTTGTTCAGCAAGACAGGAATGGCACTATCATGATGATTTGTCAAAACCAATC 143
DB 150 AACTTGTTCAGCAAGACAGGAATGGCACTATCATGATGATTTGTCAAAACCAATC 204
RESULT 5
AF282263 2994 bp DNA linear PLN 02-JUL-2001
LOCUS
DEFINITION
partial cde.
ACCESSION
AF282263
VERSION
AF282263.1 GI:14582466
KEYWORDS
ORGANISM
Phaseolus vulgaris
Phaseolus vulgaris
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 1791)
Hegeman, C.E., Good, L.L. and Grabau, E.A.
Expression of D-myo-inositol-3-phosphate synthase in soybean.
Implications for phytic acid biosynthesis
Plant Physiol. 125 (4), 1941-1948 (2001)
21196082
11299373
2 (bases 1 to 1791)
Hegeman, C.E., Good, L.L. and Grabau, E.A.
Direct Submision
Submitted (07-AUG-2000) Plant Pathology, Physiology, and Weed
Science, Virginia Tech, Fralin Biotechnology Center, Blacksburg, VA
24060, USA
Location/Qualifiers
1. 1791
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
/dev_stage="Immature cotyledons"
1. 1791
/gene="MIPS"
62. 1594
/gene="MIPS"
/note="isomerase"
/codon_start=1
/product="myo-inositol-3-phosphate synthase"
/protein_id="AAK49896.1"
/db_xref="GI:13936691"
/translation="MTENPKYKSEPNVYKSEIYSVNYETTELVHNRNGAYOMI
KPKSVYEFKNTNHPKGLVNLVGNNGNSLTGVIANRSGISWATDKIQANFY
GSLTQASAIRVSGFGEETIAPFKSLIPNVNDDIVFGWDSNNMLADAMRAVFD
IDLQKOLRPYMSWPLPGIYDPDPIANQBERANNVIGTKEQOVQIIXDKIKEKA
ATKVDKVVNLMTANTERYSNLVGNDTTEMLAADNRBAEISPTLPAICVENV
PINGSPOINTPVPGLIDLAIRNTLIGDDPFGSGOTKMSVAVDFLVGAGIKPTSV
YHLLGNNDGMLNSAPOTPRSKISNSNVDDMVNSALLYERGEHPDHVYVYKPYV
GSKRAMDEYTSIEITMGSKSTVYLNHTCDSLLAPITLIDVLAEISRTIFKANE
GKFHSFHPVATLITLTKAPLVPGTIPGNALSKOPAMLENTIRACVGLAPBNMILE
YK"

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Phaseolus.
1 (bases 1 to 2994)
Johnson, M.D., Lackey, K.H. and Pope, P.M.
Subcellular localization of 1L-myo-inositol-1-phosphate synthase
Unpublished
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
mrna
CDS
ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 2e-27;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1093 GATATGTCACAGCAAGATGCCATCTCTATGAGCCTGGGACATCCGACCATGTTGTT 1152
DB 2351 GATATGTCACAGCAAGATGCCATCTCTATGAGCCTGGGACATCCGACCATGTTGTT 2410
QY 1153 GTTATTAAAG 1161
DB 2411 GTTATTAAAG 2419
RESULT 6
AC125389/c 143012 bp DNA linear PLN 15-JUN-2004
LOCUS
DEFINITION
Medicago truncatula clone mth2-12a18, complete sequence.
ACCESSION
AC125389
VERSION
AC125389.32 GI:48717535
KEYWORDS
HTG:
SOURCE
ORGANISM
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.
1 (bases 1 to 143012)
Shauli, S., Lin, S., Dixon, R., May, G., Summer, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Medicago truncatula BAC Clone mth2-12a18
Unpublished
REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 143012)
Shauli, S., Lin, S., Dixon, R., May, G., Summer, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submision
Submitted (25-JUN-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

REFERENCE
AUTHORS
TITLE
JOURNAL
OK 73019, USA
3 (bases 1 to 143012)
Shall, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (20-MAY-2004) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 143012)
Shall, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (15-JUN-2004) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
On Jun 15, 2004 this sequence version replaced gi:47524030.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University of Oklahoma
Center code: UOKNOR

FEATURES
source
1..143012
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-12a18"
/clone_lib="Medicago truncatula BAC library H2"
ORIGIN
Query Match 3.4%; Score 52; DB 8; Length 143012;
Best Local Similarity 100.0%; Pred. No. 1.2e-17;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1072 TCCAGAGCAAGCTGTGACATATGTCACAGCAAGCATCTCTATG 1123
|||||
Db 14603 TCCAAGAGCAAGCTGTGACATATGTCACAGCAAGCATCTCTATG 14552
RESULT 7
AP006418 62713 bp DNA linear PLN 22-JUL-2003
LOCUS
DEFINITION Lotus corniculatus var. japonicus genomic DNA, chromosome 6,
clone:ljt23A02, TM0306, complete sequence.
ACCESSION
AP006418.1 GI:31581049
KEYWORDS
HTG.
SOURCE
Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM
Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteeae;
Lotus.
REFERENCE
1 Asamizu, E., Kato, T., Sato, S., Nakamura, Y., Kaneko, T. and Tabata, S.
AUTHORS
TITLE
JOURNAL
STRUCTURAL ANALYSIS OF A LOTUS JAPONICUS GENOME. IV. SEQUENCE
FEATURES AND MAPPING OF SEVENY-THREE TAC CLONES WHICH COVER THE
7.5 Mb REGIONS OF THE GENOME
DNA REE. (2003) In press
2 (bases 1 to 62713)
Sato, S.
DIRECT SUBMISSION
Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute,
Department of Plant Gene Research, 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba 292-0818, Japan (E-mail: ssato@kazusa.or.jp,
URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex. 2377),
Fax: 81-438-52-3934)
FEATURES
source
1..62713
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"

/chromosome="6"
/clone="ljt23A02"
/clone_lib="ljt library"
/note="TAC clone: TM0306-synonym: Lotus japonicus"
ORIGIN
Query Match 3.3%; Score 50; DB 8; Length 62713;
Best Local Similarity 100.0%; Pred. No. 1.8e-16;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 484 GATTTCGAGAGCAGTTCAGGCTTACATGGAATCCATGTTCCATCCC 533
|||||
Db 29198 GATTTCGAGAGCAGTTCAGGCTTACATGGAATCCATGTTCCATCCC 29247
RESULT 8
AY005128 1391 bp mRNA linear PLN 30-JAN-2001
LOCUS
DEFINITION Actinidia arguta myo-inositol-1-phosphate synthase mRNA, complete
cda.
ACCESSION
AY005128
AY005128.1 GI:12597482
KEYWORDS
SOURCE
ORGANISM
Actinidia arguta
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; Ericales; Actinidiaceae; Actinidia.
REFERENCE
1 Klages, K., Fitzgerald, A. and Moodie, M.
AUTHORS
TITLE
JOURNAL
Unpublished
2 (bases 1 to 1391)
Klages, K., Fitzgerald, A. and Moodie, M.
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (18-JUL-2000) Hortresearch, 120 Mt Albert Rd., Auckland,
New Zealand
FEATURES
source
1..1391
/organism="Actinidia arguta"
/mol_type="mRNA"
/db_xref="taxon:64478"
/tissue_type="young mature leaf"
1..1095
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAP97409.1"
/translation="MNLADAVARARVFDIDIQQLRPYMSMFLPGIYEPDPIAANO
GSRANNVYKSGKKEQVQDIKDIREFKKNKVDVVVLTMTANTERSVNTVGLNDTME
NLFPASLVKNESEIIPSTLYAACVLENI PFINGSPQNTFYGLIDLAIRNCLIGDD
FKSQOTKMSKLVLPFVAGIKPTISYVNLGNDQNLISA POTFSKESKANNVD
DMVSSNATLYVRGGRPHVVIKYVPVAVGSKRAMDYMSEIFMGKNNTVLTNHTCED
SLAAPITLIDVLAELSTRQLKADBERKSHSPVATILSYLTAKPLVPRGPVYN
ALSKORAMLENTLACIGLSPENNMLEYK"
ORIGIN
Query Match 2.9%; Score 45; DB 8; Length 1391;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 900 TTTGATTGTCGAGATGACTTCAAGAGTGTGACAGCCAAATGAA 944
|||||
Db 462 TTTGATTGTCGAGATGACTTCAAGAGTGTGACAGCCAAATGAA 506
RESULT 9
AP006419/c 119858 bp DNA linear PLN 22-JUL-2003
LOCUS
DEFINITION Lotus corniculatus var. japonicus genomic DNA, chromosome 4,
clone:ljt3018, TM0307, complete sequence.
ACCESSION
AP006419
AP006419.1 GI:31581050

KEYWORDS
HTG.
Lotus corniculatus var. japonicus (Lotus japonicus)

SOURCE
Lotus corniculatus var. japonicus

ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.

REFERENCE
1
Asamizu, E., Kato, T., Sato, S., Nakamura, Y., Kaneko, T. and Tabata, S. Structural Analysis of a Lotus japonicus Genome. IV. Sequence Features and Mapping of seventy-three TAC clones which cover the 7.5 Mb Regions of the Genome
DNA Res. (2003) In press

JOURNAL
2 (bases 1 to 119858)

AUTHORS
Sato, S.

TITLE
Direct Submission

REFERENCE
Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute, Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 252-0818, Japan (E-mail: ssato@kazusa.or.jp).
URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex. 2337), Fax: 81-438-52-3934

FEATURES
source
Location/Qualifiers
1..119858
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="4"
/clone="LJT43018"
/clone_11b="LJT 1library"
/note="TAC clone: TW0307-synonym: Lotus japonicus"

ORIGIN
Query Match 2.7%; Score 41; DB 8; Length 119858;
Best Local Similarity 100.0%; Pred. No. 2.8e-11;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
502 AGGCTTACATGGAATCCATGTTCCATCCCGGAAATCTA 542
|||||
Db 47386 AGGCTTACATGGAATCCATGTTCCATCCCGGAAATCTA 47346
|||||

RESULT 10
AB090886 1372 bp mRNA linear PLN 04-MAR-2003
LOCUS
DEFINITION
Aster tripolium mRNA for myo-inositol-1-phosphate synthase, partial cds.
AB090886
AB090886.1 GI:28804514
Aster tripolium
Aster tripolium
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Astereae; Aster.
1
Takeda, M., Uno, Y., Kanuchi, M. and Inagaki, N.
Analyze of nine cDNAs for salt-inducible genes in the halophyte sea aster
Unpublished
2 (bases 1 to 1372)
Takeda, M., Uno, Y., Kanuchi, M. and Inagaki, N.
Direct Submission
Submitted (27-AUG-2002) Migiwa Takeda, The Graduate School of Science and Technology, Kobe University, Division of Environmental Science, Rokkoudaiyuu 1-1, Kobe-shi Nada-ku 657-8501, Japan
(E-mail: 991d860n@y01.kobe-u.ac.jp, Tel: 81-78-803-5832, Fax: 81-78-803-5832)

FEATURES
source
Location/Qualifiers
1..1372
/organism="Aster tripolium"
/mol_type="mRNA"

CDS
/db_xref="taxon:74787"
<1..1230
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="BAC57963.1"
/db_xref="GI:28804515"

ORIGIN
Query Match 2.1%; Score 32; DB 8; Length 1372;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
502 AGGCTTACATGGAATCCATGTTCCATCC 533
|||||
Db 202 AGGCTTACATGGAATCCATGTTCCATCC 233
|||||

RESULT 11
CPINOIG 1978 bp DNA linear PLN 26-JUN-1995
LOCUS
DEFINITION
C. paradiisi (Macf) INOI gene.
ACCESSION
232632
232632.1 GI:602564
KEYWORDS
INOI gene.
SOURCE
Citrus x paradiisi
Citrus x paradiisi
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Sapindales; Rutaceae; Citrus.
1 (bases 1 to 1978)
Holland, D.
Direct Submission
Submitted (08-APR-1994) Doron Holland, Fruit Tree Breeding and Genetics, Volcani Research, Organisation Center, Bet Dagan, 50250, Israel

FEATURES
source
Location/Qualifiers
1..1978
/organism="Citrus x paradiisi"
/mol_type="genomic DNA"
/db_xref="taxon:37656"
/clone="c38"
/cissue_type="fully expanded leaves"
/dev_stage="juvenile plants"
188..1711
/gene="INOI"
188..1711
/gene="INOI"
/codon_start=1
/protein_id="CAA83565.1"
/db_xref="GI:602565"
/db_xref="GOA:P42802"
/db_xref="UniProt/Swiss-Prot:P42802"
/translation="MFLENPKYSEPNVYKTHREHSVYDYETTELVEHNRNTQYMI
RKTYREKTDVHVPRKGVMLVGGKNGSLTGTGVIANRGTICWATKDVQOANYF
GSLTQASAIRVGSVYGEIIVAFKSIILNVNDDIVFGWMDISDNLADNARAVFD
IDLQKRLPYMESWVPLPGIYDPPRIANQGRANVYKGTKEQMLQIIKDIRFKE
KKRVRRVVLMTANERSNVVGLNDVSESLASLDKKEAISSTLYAIACTVLENI
PINSNPNTFVPGIDLAIRNCLIGDDDFSGQTKKSVLVDFLVAGAIKPTVIS
VNHGNNQNMNLAPOTRSEKISVNVDDMVSNVPMGLVNTPRPWIKYVPAVAI
BRANDYSEIFPMGSKTIVLHNTEDSLAIPILDLVYLAEISTRIQLKRGEGKF
HSFHPVATILSLTAPLVPPTPVNALSQRALENILRACVGLADENNMTLEYK"

ORIGIN
Query Match 2.1%; Score 32; DB 8; Length 1978;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 ATGCTTGTTGGGTGAACAACGCGCTC 224
|||||
Db 380 ATGCTTGTTGGGTGAACAACGCGCTC 411

RESULT 12
MCU32511 2053 bp mRNA linear PLN 11-JUL-1996
LOCUS Mesembryanthemum crystallinum myo-inositol-1-phosphate synthase
DEFINITION
ACCESSION U32511
VERSION U32511.1 GI:975887
KEYWORDS
SOURCE
ORGANISM Mesembryanthemum crystallinum (common iceplant)
REFERENCE
AUTHORS Mesembryanthemum crystallinum
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Aizoaceae; Mesembryanthemum.
1 (bases 1 to 2053)
Ishitani,M., Majumder,A.L., Bornhouser,A., Michalowski,C.B.,
Jensen,R.G. and Bohner,H.U.
Coordinate transcriptional induction of myo-inositol metabolism
during environmental stress
Plant J. 9 (4), 537-548 (1996)
96208959
8624516
2 (bases 1 to 2053)
Ishitani,M., Majumder,A.L., Bornhouser,A., Michalowski,C.B.,
Jensen,R.G. and Bohner,H.U.
Direct Submission
Submitted (27-JUL-1995) Hans U. Bohner, Biochemistry, University
of Arizona, Biosciences West 516, Tucson, AZ 85721, USA
FEATURES
source
Location/Qualifiers
1..2053
/organism="Mesembryanthemum crystallinum"
/mol_type="mRNA"
/db_xref="taxon:3544"
182..1720
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AA03687.1"
/db_xref="GI:975888"
/translation="MFISFVYSPVNVKXTEIEISVYVDTTELHNRKQAGYOM
IVKPTVOYHFKTDVPLKGLVGMVGNNNGSTLGVIANREGISWATKXIOAN
YRGSLTQASSTIVGSPNGEITAPFKSLPMPNPPDVVFGNDIDNALADMTARV
FDIDLOKQLRPTMEHWPPLGIDPDIANQSRANNVKIKTKQYERVKDIREF
KEKNVYDKRVVLTMTGTERYSNVVGLDTMNLASLEKSEISPSSTVALACIEE
NIPFINGSPONTFVPLIDLAIKKNSLIGDDPFKSGQTKMSVLVDFVAGIKPTSI
VYNHGGNDGNNLSAPOTFRSEKISKSVDDPWSANGIILEPGEHDPHVAVIKVP
VYDSCRAMDEYTSIEFMGCTNTIVMHNHCESLLAAPILIDVLVLAELSRIOKAE
EBDKRSHFPAVTITLSYLTKAPLVPPGTPVVALAKORAMLENIRACVGLAPENNMI
LEYK"

ORIGIN
Query Match 1.9%; Score 29; DB 8; Length 2053;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1183 AGCAAGAGCCATGCATGATGACTTC 1211
|||||
Db 1370 AGCAAGAGCCATGCATGATGACTTC 1398

RESULT 13
AY333824 1533 bp mRNA linear PLN 15-JUL-2003
LOCUS Xerophyta viscosa myo-inositol-1-phosphate synthase INO1 mRNA,
DEFINITION complete cds.
ACCESSION AY333824
VERSION AY333824.1 GI:32492881
KEYWORDS

SOURCE
ORGANISM Xerophyta viscosa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Velloziaceae; Xerophyta.
REFERENCE
AUTHORS 1 (bases 1 to 1533)
Majee,M., Majumder,A.N.L. and Munder,S.G.
TITLE Molecular characterization of Xvino1, a myo-inositol-1-phosphate
synthase from the resurrection plant Xerophyta viscosa
JOURNAL Unpublished
2 (bases 1 to 1533)
Majee,M., Majumder,A.N.L. and Munder,S.G.
AUTHORS Direct Submission
TITLE Submitted (14-JUN-2003) Molecular and Cell Biology, University of
Cape Town, Private Bag, Rondebosch, Western Cape 7701, South Africa
FEATURES
source
Location/Qualifiers
1..1533
/organism="Xerophyta viscosa"
/mol_type="mRNA"
/db_xref="taxon:90708"
1..1533
/note="Xvino1"
/codon_start=1
/product="myo-inositol-1-phosphate synthase INO1"
/protein_id="AP8553.1"
/db_xref="GI:32492882"
/translation="MFIESPVSPPHYRGAAEISDYQYDTTELVHSHDASRWV
RPSVRYNFRFTTVPLKGLVLMVGNGNSTTLRAGVIANREGISWATKQVQOANY
GLTQASSTIRVGSYNGEITAPFKSLPMPNPPDVVFGNDIDNALADMTARVLD
IDLOKQLRPTMEHWPPLGIDPDIANQSRANNVKIKTKQYERVKDIREF
KSKYDKRVVLTMTGTERYSNVCGVGLNTMNLASVDNKAIEISPSSTVALACMEGI
PFINGSPONTFVPLIDLAIKKNSLIGDDPFKSGQTKMSVLVDFVAGIKPTSIYS
VYNHGGNDGNNLSAPOTFRSEKISKSVDDPWSANGIILEPGEHDPHVAVIKVPY
GDSKRAMDEYTSIEFMGCTNTIVMHNHCESLLAAPILIDVLVLAELSTRIOKAE
EKPSHFPAVTITLSYLTKAPLVPPGTPVVALAKORAMLENIRACVGLAPENNMI
LEYK"

ORIGIN
Query Match 1.7%; Score 26; DB 8; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1105 AGCAATGCCATCCTATGACCTGG 1130
|||||
Db 1105 AGCAATGCCATCCTATGACCTGG 1130

RESULT 14
AY096554 1564 bp mRNA linear PLN 18-SEP-2002
LOCUS Arabidopsis thaliana putative myo-inositol-1-phosphate synthase
DEFINITION (AF591017) mRNA, complete cds.
ACCESSION AY096554
VERSION AY096554.1 GI:20465488
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1564)
Yanada,K., Bann,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamliya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Meyers,W.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Arabidopsis Open Reading Frame (ORF) Clones
JOURNAL Unpublished
2 (bases 1 to 1564)
Yanada,K., Bann,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamliya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Meyers,W.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Arabidopsis Open Reading Frame (ORF) Clones
JOURNAL Unpublished
2 (bases 1 to 1564)
Yanada,K., Bann,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,

TITLE
JOURNAL
COMMENT

Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PEGC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.C., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PEGC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PEGC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.
Location/Qualifiers
1..1564
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U18590"
/ecotype="Columbia"
/note="this clone is in pUNI 51."
1..1564
/gene="At5g10170"
1..1533
/gene="At5g10170"
/codon_start=1
/evidence=experimental
/product="putative myo-inositol-1-phosphate synthase"
/protein_id="AA020204.1"
/db_xref="GI:20465489"
/translation="MFIESPKVESPVVXTENSHSVYDTTELVHKNKGAFOKTVKRTVYVEKTDTHVPLKGLVLMVGNNGNSTITAGVIANREGISWATKEKYQOANYF GSLTQASIRVSPNGEEIYAPPKSLPMVNPBEIVFGWDISDMNLADAMARAVLD IDLOKQMRPFMEHWPPLPGIPDPDIANQGSANVVIKGTKKQOLEQVIXIDIREPKE KKRVDPVLTANTERYSNVVGLNDITENLMSLIDKDEAEISPTLYAICVLENV PFINGSPONTFWGLIELAIKNKCLIGDDPKSGOTKMSVLYDFVAGIIPKTSIVS VNHIGNNDGMNLISAPOTFSKESKSNVVDWVNSGILYEPGEHPDHVVYI KYVPCV GBSKRAMDEYTSIEIPFMGKRTTYMHTCDSLALAPIIDLVLLAELTRIQMSENE GKFSHPFVATLLSTYSKAPLVPGTGVVNALSKQAPMLNVLACVGLAPENNMLE YK"
1534..1564
/gene="At5g10170"

3' UTR
ORIGIN

Query Match 1.7%; Score 26; DB 8; Length 1564;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 913 GATGACTCAGAGTGTGACGACCA 938
DB 913 GATGACTTCAAGAGTGTGACGACCA 938

RESULT 15
AX054630

LOCUS AX054630 1781 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 1 from Patent WO0073473.
ACCESSION AX054630
VERSION AX054630.1 GI:12228190
KEYWORDS
SOURCE
ORGANISM
Brassica napus (rape)
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1
Georges, F., Hussain, A.A. and Keller, W.A.
Method for reducing phytate in canola meal using genetic manipulation involving myo-inositol 1-phosphate synthase gene Patent: WO 0073473-A 1 07-DEC-2000;
NATIONAL RESEARCH COUNCIL CANADA (CA)
FEATURES
source
1..1781
/organism="Brassica napus"
/mol_type="unassigned DNA"
/db_xref="taxon:3708"

ORIGIN

Query Match 1.7%; Score 26; DB 6; Length 1781;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 907 GGTTGAGATGACTTCAGAGTGATCA 932
DB 963 GGTTGAGATGACTTCAGAGTGATCA 988

Search completed: June 8, 2005, 11:26:38
Job time : 6690.71 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:47:21 ; Search time 842.149 Seconds
(without alignments)
10775.958 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533
Sequence: 1 atgttcacgcgagaatttaa.....tgattctcgagacagtgga 1533

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1533	100.0	1533	12	ADQ14500	Adq14500 Mutant so
2	1533	100.0	1533	13	AD882001	Ad882001 Soybean m
3	1482	96.7	1533	12	ADQ14504	Adq14504 Wild type
4	1482	96.7	1533	12	ADQ14502	Adq14502 Mutant so
5	1482	96.7	1533	13	AD882005	Ad882005 Soybean m
6	1482	96.7	1533	13	AD882003	Ad882003 Soybean m
7	146	9.5	1533	2	AAV62443	AAV62443 Soybean m
8	146	9.5	1533	12	ADQ14498	Adq14498 Mutant so
9	146	9.5	1533	12	ADQ14494	Adq14494 Mutant so
10	146	9.5	1533	13	AD881999	Ad881999 Soybean m
11	146	9.5	1533	13	AD881993	Ad881993 Soybean m
12	146	9.5	1533	13	AD881997	Ad881997 Soybean m
13	146	9.5	1760	12	ADQ14490	Adq14490 Wild type
14	146	9.5	1760	12	AAV62440	AAV62440 Soybean w
15	26	1.7	377	2	ADP93370	Adp93370 Cotton ex
16	26	1.7	532	13	ACN61638	Acn61638 Cotton gy
17	26	1.7	533	13	ACN62170	Acn62170 Cotton gy
18	26	1.7	543	13	ACN61870	Acn61870 Cotton gy
19	26	1.7	566	13	ACN57874	Acn57874 Cotton gy
20	26	1.7	618	13	ACN62990	Acn62990 Cotton ca

21	26	1.7	1781	4	AC87643	Aac87643 Brassica
22	26	1.7	1950	2	AAV90402	AAV90402 Nicotiana
23	23	1.5	35	2	AAV62441	AAV62441 Soybean m
24	23	1.5	35	12	ADQ14492	Adq14492 Soybean m
25	23	1.5	345	12	ADP94497	Adp94497 Cotton ex
26	23	1.5	452	12	ADP94561	Adp94561 Cotton ex
27	23	1.5	557	13	ACN58059	Acn58059 Cotton gy
28	23	1.5	601	13	ACN60643	Acn60643 Cotton gy
29	23	1.5	605	13	ACN60421	Acn60421 Cotton gy
30	23	1.5	1536	12	ADP43918	Adp43918 P. coatc
31	23	1.5	1578	2	AAV91656	AAV91656 Inositol-
32	23	1.5	1578	2	AAV65400	AAV65400 Inducible
33	23	1.5	1665	2	AAV09006	AAV09006 Maize myo
34	23	1.5	1931	2	AAV24407	AAV24407 Maize myo
35	23	1.5	3546	2	AAV24410	AAV24410 Maize myo
36	23	1.5	3546	2	AAV24411	AAV24411 Maize myo
37	21	1.4	35	13	AD881995	Ad881995 PCR prime
38	21	1.4	39	12	ADQ14493	Adq14493 Soybean m
39	21	1.4	39	12	ADQ14493	Adq14493 Soybean m
40	20	1.3	291	12	ADP93783	Adp93783 Cotton ex
41	20	1.3	388	13	ACN52468	Acn52468 Cotton an
42	20	1.3	389	13	ACN51365	Acn51365 Cotton an
43	20	1.3	449	12	ADP91072	Adp91072 Cotton ex
44	20	1.3	506	13	ACN57963	Acn57963 Cotton gy
45	20	1.3	517	13	ACN63079	Acn63079 Cotton ca

ALIGNMENTS

RESULT 1	ADQ14500	ADQ14500 standard; cDNA, 1533 BP.
XX	XX	ADQ14500;
AC	XX	ADQ14500;
XX	XX	23-SEP-2004 (first entry)
DT	XX	Mutant soybean myo-inositol 1-phosphate synthase cDNA #3.
XX	XX	Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW	XX	myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KM	XX	inorganic phosphate; mutant.
XX	XX	Glycine max.
OS	XX	Synthetic.
XX	XX	Key
FH	XX	Location/Qualifiers
FT	XX	1.1533
FT	XX	/tag=a
FT	XX	/product="Mutant soybean myo-inositol 1-phosphate
FT	XX	synthase #3"
XX	XX	US2004128713-A1.
PN	XX	01-JUL-2004.
PD	XX	21-NOV-2003; 2003US-00718952.
XX	XX	08-APR-1997; 97US-00835751.
PR	XX	07-APR-1998; 98WO-US006822.
PR	XX	26-APR-1999; 99US-00299315.
PR	XX	11-MAR-2002; 2002US-00025003.
XX	XX	(HITZ/) HITZ W D.
PA	XX	(SEBA/) SEBASTIAN S A.
PA	XX	(GRAC/) GRACE D J.
PA	XX	(STRE/) STREIT L G.
XX	XX	Hitz WD, Sebastian SA, Grace DJ, Streit LG;
PI	XX	WPI; 2004-53335/51.
DR	XX	P-PsDB; ADQ14501.

XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

PS Claim 10; SEQ ID NO 11; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

Query Match 100.0%; Score 1533; DB 12; Length 1533;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANGTCATCGAAGATTTTAAGTAGAGAGTCTTAATGTAACCGAGACTGAGATT 60
DB 1 AAGTCATCGAAGATTTTAAGTAGAGAGTCTTAATGTAACCGAGACTGAGATT 60
QY 61 CAGTCGCTGTACACTACGAAACACCGAACTTGTTCAGAGAGAGAGATGCACTAT 120
DB 61 CAGTCGCTGTACACTACGAAACACCGAACTTGTTCAGAGAGAGAGATGCACTAT 120
QY 121 CAGTGGATTTGCAAAACCCAAATCCGTCACATTAATTAACCAACCCATGTTCCA 180
DB 121 CAGTGGATTTGCAAAACCCAAATCCGTCACATTAATTAACCAACCCATGTTCCA 180
QY 121 CAGTGGATTTGCAAAACCCAAATCCGTCACATTAATTAACCAACCCATGTTCCA 180
DB 121 CAGTGGATTTGCAAAACCCAAATCCGTCACATTAATTAACCAACCCATGTTCCA 180
QY 181 AATTTGGGGGATGATCGTTGGGGTGGGATGGAACAACGGCTCAACCGGTGGT 240
DB 181 AATTTGGGGGATGATCGTTGGGGTGGGATGGAACAACGGCTCAACCGGTGGT 240
QY 241 GTTATTTGCTAACAGAGAGACATTTTCATGGCTACAAAGAGACAAATTCACCAAT 300
DB 241 GTTATTTGCTAACAGAGAGACATTTTCATGGCTACAAAGAGACAAATTCACCAAT 300
QY 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCAGATTTGATCTTCCAGGAGAGAA 360
DB 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCAGATTTGATCTTCCAGGAGAGAA 360
QY 361 ACTATGCCCCCATTAAGAGTGTGCTTCCATAGTTAATCCAGAGACATTTGTTGGG 420
DB 361 ACTATGCCCCCATTAAGAGTGTGCTTCCATAGTTAATCCAGAGACATTTGTTGGG 420
QY 421 GGATGGGATATCAGCAACATGACCTGCTGATGTCATGGCCAGAGGCAAAAGTGTGAC 480
DB 421 GGATGGGATATCAGCAACATGACCTGCTGATGTCATGGCCAGAGGCAAAAGTGTGAC 480
QY 481 ATCGATTTGCAAGACAGTTGAGGCTTACATGAAATCATGTTCCACTCCCGGAATC 540
DB 481 ATCGATTTGCAAGACAGTTGAGGCTTACATGAAATCATGTTCCACTCCCGGAATC 540
QY 541 TACGACCCGGATTTTATGCTGCTCAACCAAGAGGAGCGTCCCAACACGTGATTAAGGCG 600
DB 541 TACGACCCGGATTTTATGCTGCTCAACCAAGAGGAGCGTCCCAACACGTGATTAAGGCG 600

QY 601 ACAAAGCAAGAGCAAGTTACAGCAATATCATCAAGACATCAAGCGCTTTAAGAAACCAAC 660
DB 601 ACAAAGCAAGAGCAAGTTACAGCAATATCATCAAGACATCAAGCGCTTTAAGAAACCAAC 660
QY 661 AAAGTGCACAAAGTGTGCTGCTGATGCTGCACTGCAACAGAGAGATTAAGCAATTTGGTT 720
DB 661 AAAGTGCACAAAGTGTGCTGCTGATGCTGCACTGCAACAGAGAGATTAAGCAATTTGGTT 720
QY 721 GTAGGCTTTATGACACCATGAGAAATCTTGGCTGCTGTCAGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTTATGACACCATGAGAAATCTTGGCTGCTGTCAGACAGAAATGAGGCTGAG 780
QY 781 ATTTCCTCTTCACTTGTATGATGCTGCTGATGATGAAATGTTCTTTCAATTAAT 840
DB 781 ATTTCCTCTTCACTTGTATGATGCTGCTGATGATGAAATGTTCTTTCAATTAAT 840
QY 841 GGAAGCCTTCAGAACATTTTGTACAGAGGCTGATGATCTTGGCATGCGAGAACACT 900
DB 841 GGAAGCCTTCAGAACATTTTGTACAGAGGCTGATGATCTTGGCATGCGAGAACACT 900
QY 901 TTGATTTGGTGAAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGAT 960
DB 901 TTGATTTGGTGAAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGAT 960
QY 961 TTTCTTGGGGGCTGATATCAAGCCCAATCTATTAAGTTAATCAACCATCTGGGAAAC 1020
DB 961 TTTCTTGGGGGCTGATATCAAGCCCAATCTATTAAGTTAATCAACCATCTGGGAAAC 1020
QY 1021 AATGATGATGAATCTTGGCTTCCAAACCTTCCGCTCCAGAAATCTCCAAAGGC 1080
DB 1021 AATGATGATGAATCTTGGCTTCCAAACCTTCCGCTCCAGAAATCTCCAAAGGC 1080
QY 1081 AAGCTTTGTCAGATATGATGTCACAGCAATGCCATCTCTATGAGCCTGTGAAATCTCC 1140
DB 1081 AAGCTTTGTCAGATATGATGTCACAGCAATGCCATCTCTATGAGCCTGTGAAATCTCC 1140
QY 1141 GACCATGTTGTTTATTAAGTATGCTCTTACGTAAGGGATATGCAAGAGACCATGGAT 1200
DB 1141 GACCATGTTGTTTATTAAGTATGCTCTTACGTAAGGGATATGCAAGAGACCATGGAT 1200
QY 1201 GAGTACATTTGACAGATTTTCAATGAGTGGGAAAGAACACATTTGTTGACAAACATATG 1260
DB 1201 GAGTACATTTGACAGATTTTCAATGAGTGGGAAAGAACACATTTGTTGACAAACATATG 1260
QY 1261 GAGGATTCCTTTTAACTGCTCTTATTAATCTTGAATCTGCTTCTTGTGCTGAGCTGAGC 1320
DB 1261 GAGGATTCCTTTTAACTGCTCTTATTAATCTTGAATCTGCTTCTTGTGCTGAGCTGAGC 1320
QY 1321 ACTGAATCAAGTTTAAAGCTGAAAAATGAGGAAAAATTCACATCTTCCAGCTGAGTGGCT 1380
DB 1321 ACTGAATCAAGTTTAAAGCTGAAAAATGAGGAAAAATTCACATCTTCCAGCTGAGTGGCT 1380
QY 1381 ACCATTTCTACGATATCTGACCAAGGCTCTCTGTTCCACCGGGTACACAGTGGTGAAT 1440
DB 1381 ACCATTTCTACGATATCTGACCAAGGCTCTCTGTTCCACCGGGTACACAGTGGTGAAT 1440
QY 1441 GCATTTGCAAAAGCAGCGTGAATGCTGGAAGAACTAATGAGAGGCTTGTGTTGATTTGGCC 1500
DB 1441 GCATTTGCAAAAGCAGCGTGAATGCTGGAAGAACTAATGAGAGGCTTGTGTTGATTTGGCC 1500
QY 1501 CCAGGAATTAACATGATTTCTGAGTACAAAGTGA 1533
DB 1501 CCAGGAATTAACATGATTTCTGAGTACAAAGTGA 1533

RESULT 2

ADSS82001

ID ADSS82001 standard; cDNA; 1533 BP.

XX ADSS82001;

XX

DT 18-NOV-2004 (first entry)

XX

DE Soybean myo-inositol 1-phosphate synthase cDNA mutant #2.
 XX
 KM Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
 XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
 OS Glycine max; line 29010CP01.
 XX
 XX Synthetic.
 FT Key Location/Qualifiers
 FT CDS 1..1533
 FT /tag= a
 FT /product= "myo-inositol 1-phosphate synthase"
 FT mutation replace(260,G)
 FT /tag= b
 PN US2003074685-A1.
 PD 17-APR-2003.
 XX
 XX 11-MAR-2002; 2002US-00025003.
 PF 08-APR-1997; 97US-00835751.
 XX
 PR 07-APR-1998; 98WO-US006822.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PI Hitz WD, Sebastian SA;
 XX
 XX WPI; 2004-639957/62.
 DR P-PSDB; ADS82002.
 XX
 PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
 PT phytic acid and inorganic phosphate content of soybean seeds.
 PS
 XX
 XX Claim 8; SEQ ID NO 11; 34pp; English.
 XX
 XX The invention relates to an isolated nucleic acid fragment encoding a
 CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 CC phosphate synthase having decreasing capacity for the synthesis for myo-
 CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
 CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 CC complement, subfragment or the complement of the subfragment, operably
 CC linked to suitable regulatory sequences, where expression of the chimeric
 CC gene results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
 CC acid content of less than 17 micromol/g), a seed content of raffinose plus
 CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
 CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype
 CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homozygous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with a
 CC heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence encodes a mutant myo-
 CC inositol 1-phosphate synthase.
 CC
 XX
 XX Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 1533; DB 13; Length 1533;
 Best Local Similarity 100.0%; Pred. No. 0;

	Matches 1533;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	ATGTTTCATGAAATTTTAAGTAGAGTCTTAATGTGAAGTACCCGAGCTGAGATT							60
DB	1	ATGTTTCATGAAATTTTAAGTAGAGTCTTAATGTGAAGTACCCGAGCTGAGATT							60
QY	61	CAGTCCGCTGTAACATGAGAAACACCCGAGCTGTTTCAAGAAACGAAATGGACCTTAT							120
DB	61	CAGTCCGCTGTAACATGAGAAACACCCGAGCTGTTTCAAGAAACGAAATGGACCTTAT							120
QY	121	CAGTGGATTGTCAAAACCCAAATCCGTCACATCAATTTAAACCAACCAACCATGTTCCA							180
DB	121	CAGTGGATTGTCAAAACCCAAATCCGTCACATCAATTTAAACCAACCAACCATGTTCCA							180
QY	181	AAATTGGGGGTATATGCTTGTGGGTGGGATGAAACACGCTCTACCCCTACCCGGTGT							240
DB	181	AAATTGGGGGTATATGCTTGTGGGTGGGATGAAACACGCTCTACCCCTACCCGGTGT							240
QY	241	GTTATTTGCTTAAGAGAGAGACATTTTCATAGGCTTACAAAGACAAAGATTCAAGCCAT							300
DB	241	GTTATTTGCTTAAGAGAGAGACATTTTCATAGGCTTACAAAGACAAAGATTCAAGCCAT							300
QY	301	TACTTTGGCTCCCTCAGCCCAAGCTCAGCTATTGAGTTGATTCCTTCCAGGAGAGAA							360
DB	301	TACTTTGGCTCCCTCAGCCCAAGCTCAGCTATTGAGTTGATTCCTTCCAGGAGAGAA							360
QY	361	ATCTATGCCCATTCAGAGTCTGCTTCCATGTGTTATCTTGACGACATTTGTGTTGGG							420
DB	361	ATCTATGCCCATTCAGAGTCTGCTTCCATGTGTTATCTTGACGACATTTGTGTTGGG							420
QY	421	GGATGGGAAATGAGAAACATGAACTGGCTGATGCTGATGCTGCAAGGAGGAAAGGTTTGC							480
DB	421	GGATGGGAAATGAGAAACATGAACTGGCTGATGCTGATGCTGCAAGGAGGAAAGGTTTGC							480
QY	481	ATCGATTGTCAGAAACAGTTGAGGCTTACATGATCATGATTCCTCCCTCCCGAATC							540
DB	481	ATCGATTGTCAGAAACAGTTGAGGCTTACATGATCATGATTCCTCCCTCCCGAATC							540
QY	541	TACGACCCGGATTTTCATGCTGCTCCAAACGAGAGAGCTGCAACAAACGTATTAAGGCG							600
DB	541	TACGACCCGGATTTTCATGCTGCTCCAAACGAGAGAGCTGCAACAAACGTATTAAGGCG							600
QY	601	AAAAAGCAAGACCAAGTTCAGAAATCATCAAGATCAAGCCCTTTAAAGAAAGCAAC							660
DB	601	AAAAAGCAAGACCAAGTTCAGAAATCATCAAGATCAAGCCCTTTAAAGAAAGCAAC							660
QY	661	AAAGTGACCAAGGTTGCTCTGTGACTGCAACACAGAGGATATGCAATTTGGTT							720
DB	661	AAAGTGACCAAGGTTGCTCTGTGACTGCAACACAGAGGATATGCAATTTGGTT							720
QY	721	GTAGGCTTTAATGACACATGAGAAATCTTTGCTGTGTGAGACAAATAGAGCTGAG							780
DB	721	GTAGGCTTTAATGACACATGAGAAATCTTTGCTGTGTGAGACAAATAGAGCTGAG							780
QY	781	ATTTCTCTTCCACCTTGTATGCCATTGCTGTGTGATGAAATATTTCTTTCAATAT							840
DB	781	ATTTCTCTTCCACCTTGTATGCCATTGCTGTGTGATGAAATATTTCTTTCAATAT							840
QY	841	GGAGGCCCTCAGAACATTTTATACAGGCTGATTTGATCTTGCCATGCGAGAACT							900
DB	841	GGAGGCCCTCAGAACATTTTATACAGGCTGATTTGATCTTGCCATGCGAGAACT							900
QY	901	TTGATTTGAGATGACTTCAAGAGTGTGACACCAAAATGTAATCTGTGTGTTGAT							960
DB	901	TTGATTTGAGATGACTTCAAGAGTGTGACACCAAAATGTAATCTGTGTGTTGAT							960
QY	961	TTTCTTGGGGGCTGTGTATCAAGCAACATTAATAGTTAATCAACATCTGGGAAAC							1020
DB	961	TTTCTTGGGGGCTGTGTATCAAGCAACATTAATAGTTAATCAACATCTGGGAAAC							1020
QY	1021	AATGATGATGAATCTCTCGGCTCCAAACCTTCCGCTCAAGAAATCTCCAAAGAC							1080
DB	1021	AATGATGATGAATCTCTCGGCTCCAAACCTTCCGCTCAAGAAATCTCCAAAGAC							1080

QY 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTCTATAGCCTGTGTAACATCCC 1140
DB 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTCTATAGCCTGTGTAACATCCC 1140
QY 1141 GACCATGTTGTTGTTATTAATGATGTGCTTTCGTAAGGAGATGCAAGAGCCCATGGAT 1200
DB 1141 GACCATGTTGTTGTTATTAATGATGTGCTTTCGTAAGGAGATGCAAGAGCCCATGGAT 1200
QY 1201 GAGTACACTTCGAGATATTCATGAGGTGGAAAGAACCATGTTTGGACAACATGAT 1260
DB 1201 GAGTACACTTCGAGATATTCATGAGGTGGAAAGAACCATGTTTGGACAACATGAT 1260
QY 1261 GAGGATTCCTTTTGAAGCTGCTCTATTAATCTTGAAGCTGCTTCTTGTGAGCTGAGC 1320
DB 1261 GAGGATTCCTTTTGAAGCTGCTCTATTAATCTTGAAGCTGCTTCTTGTGAGCTGAGC 1320
QY 1321 ACTAGATTCAGTTTAAGCTGAAATGAGGAAATTCCTCACTTCCACCCAGTTGCT 1380
DB 1321 ACTAGATTCAGTTTAAGCTGAAATGAGGAAATTCCTCACTTCCACCCAGTTGCT 1380
QY 1381 ACCATTCTAGCTATCTGACCAAGGCTCTGTTCCACGGGTAACACAGTGGAT 1440
DB 1381 ACCATTCTAGCTATCTGACCAAGGCTCTGTTCCACGGGTAACACAGTGGAT 1440
QY 1441 GCATTGTCAAGCAGCTGCAATGCTGGAACCATATGAGGGCTTGTGATTTGGCC 1500
DB 1441 GCATTGTCAAGCAGCTGCAATGCTGGAACCATATGAGGGCTTGTGATTTGGCC 1500
QY 1501 CCAGAGATACATGATTTCTCGAGTACAAAGTGA 1533
DB 1501 CCAGAGATACATGATTTCTCGAGTACAAAGTGA 1533

RESULT 3
ADQ14504
ID ADQ14504 standard; cDNA; 1533 BP.

AC ADQ14504;

XX 23-SEP-2004 (first entry)

DE Wild type soybean myo-inositol 1-phosphate synthase cDNA #2.

KM Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
inorganic phosphate.

XX Glycine max.

XX Key Location/Qualifiers

FT CDS 1..1533
FT /tag="a
FT /product="Wild type soybean myo-inositol 1-phosphate
FT synthase #2"

PN US2004128713-A1.

PD 01-JUL-2004.

PF 21-NOV-2003; 2003US-00718952.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.

PR 26-APR-1999; 99US-00299315.

PR 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.

PA (STRE/) STREIT L G.

PI HITZ WD, Sebastian SA, Grace DJ, Streit LG;

XX WPI; 2004-533135/51.
DR P-PSDB; ADQ14505.

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

XX Claim 4; SEQ ID NO 15; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 96.7%; Score 1482; DB 12; Length 1533;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTTCATCGAGATTTTAAGGTAGAGAGTCTTAATGTAAGTACCGAGACTAGATT 60
DB 1 ATGTTTCATCGAGATTTTAAGGTAGAGAGTCTTAATGTAAGTACCGAGACTAGATT 60
QY 61 CAGTCGGGTCAACTAGAGAAACCGACACTGTTGTCAGAGAACAGAAATGGACCTAT 120
DB 61 CAGTCGGGTCAACTAGAGAAACCGACACTGTTGTCAGAGAACAGAAATGGACCTAT 120
QY 121 CAGTGGATTTGTCAAACCCAAATCCGTCACCTAATTAACCAATTAACCAACCATGTTCCA 180
DB 121 CAGTGGATTTGTCAAACCCAAATCCGTCACCTAATTAACCAATTAACCAACCATGTTCCA 180
QY 181 AAATTGGGGGTGATCTTGTGGGTGGGTAACCAACCGCTCACTCCCTACCGGTGCT 240
DB 181 AAATTGGGGGTGATCTTGTGGGTGGGTAACCAACCGCTCACTCCCTACCGGTGCT 240
QY 241 GTTATTGCTAACAGAGAGACATTTATGAGGCTACAAAGACCAAGTTCACCAACCAAT 300
DB 241 GTTATTGCTAACAGAGAGACATTTATGAGGCTACAAAGACCAAGTTCACCAACCAAT 300
QY 301 TACTTTGGCTCCCTCACCACCAAGCTCAGCTATTGCAAGTTGATTCCTTCAGGAGAGANA 360
DB 301 TACTTTGGCTCCCTCACCACCAAGCTCAGCTATTGCAAGTTGATTCCTTCAGGAGAGANA 360
QY 361 ATCTATGCCCATTCAGAGATGCTGCTTCAATGTTAACTCTGACGACATTTGTTGGG 420
DB 361 ATCTATGCCCATTCAGAGATGCTGCTTCAATGTTAACTCTGACGACATTTGTTGGG 420
QY 421 GGAATGGATATCAGAAATGAACCTGCTGATGCAATGAGGCAAGGAGGATGTTGAC 480
DB 421 GGAATGGATATCAGAAATGAACCTGCTGATGCAATGAGGCAAGGAGGATGTTGAC 480
QY 481 ATCGATTGCAAGAACAGATTGAGGCTTACATGATCATGATGTTCACTCCCGGAATC 540
DB 481 ATCGATTGCAAGAACAGATTGAGGCTTACATGATCATGATGTTCACTCCCGGAATC 540
QY 541 TAGACCCGGATTTTCATTGCTGCAACCAAGAGAGCGTCCCAACAGCTGATTAAGGCG 600


```

Db      1 ATGTTCAATCGAAGATTTTAAGGTAGAGAGTCTCTAATGTGAAGTACACCGAGCTGAGATT 60
Qy      61 CAGTCCGTGTACAACTACGAAACCAACCGAATCTGTGTACAGAGACAGGAAATGGCACTAT 120
Db      61 CAGTCCGTGTACAACTACGAAACCAACCGAATCTGTGTACAGAGACAGGAAATGGCACTAT 120
Qy      121 CAGTGGATTTGCAAAACCCAAATCCGTCAACCTAACCAATTTAAAAACAACACCCATGTTCCA 180
Db      121 CAGTGGATTTGCAAAACCCAAATCCGTCAACCTAACCAATTTAAAAACAACACCCATGTTCCA 180
Qy      181 AAATTGGGGGTGATGCTGTGGGTGGGGTGGGAAACAACGGCTTACCTCCACGGGTGT 240
Db      181 AAATTGGGGGTGATGCTGTGGGTGGGAAACAACGGCTTACCTCCACGGGTGT 240
Qy      241 GTTATTGCTACAGAGAGACATTTTCATGGGCTACAAAGAGACAAAGATTCAACAACCAAT 300
Db      241 GTTATTGCTACAGAGAGAGACATTTTCATGGGCTACAAAGAGACAAAGATTCAACAACCAAT 300
Qy      301 TACTTTGGCTCCCTGACCCCAAGCCTCAAGCTATTGAGATTGATCTTCCAGGGAGAGAA 360
Db      301 TACTTTGGCTCCCTGACCCCAAGCCTCAAGCTATTGAGATTGATCTTCCAGGGAGAGAA 360
Qy      361 ATCTATGCCCCCATTTCAAGAGTCTGCTTCCATGTTAAATCTTGACGACATTTGTTGGG 420
Db      361 ATCTATGCCCCCATTTCAAGAGTCTGCTTCCATGTTAAATCTTGACGACATTTGTTGGG 420
Qy      421 GGATGGGATATCAGCAACAATGAACTGGCTGATGTCATGGCCAGGGCAAAAGGTTTGAC 480
Db      421 GGATGGGATATCAGCAACAATGAACTGGCTGATGTCATGGCCAGGGCAAAAGGTTTGAC 480
Qy      481 ATGATTTTGACAGACAGTTGAGAGCCTTAATGAAATTCATGTTTCACTCCCGGAATC 540
Db      481 ATGATTTTGACAGACAGTTGAGAGCCTTAATGAAATTCATGTTTCACTCCCGGAATC 540
Qy      541 TACGACCCCGGATTTTATGCTGCTGCCAACGAGAGAGCGGCCAACAACTGTTAAAGGC 600
Db      541 TACGACCCCGGATTTTATGCTGCTGCCAACGAGAGAGCGGCCAACAACTGTTAAAGGC 600
Qy      601 ACAAGCAAGACCAAGTTCAGCAAAATCATCAAGACATCAAGGCGTTTAAAGAACCCACC 660
Db      601 ACAAGCAAGACCAAGTTCAGCAAAATCATCAAGACATCAAGGCGTTTAAAGAACCCACC 660
Qy      661 AAAGTGAACAGAGTGTGTCTCTGTGACTGCGCAACACAGAGAGTATGCAATTTGGTT 720
Db      661 AAAGTGAACAGAGTGTGTCTCTGTGACTGCGCAACACAGAGAGTATGCAATTTGGTT 720
Qy      721 GTAGGCTTTAATGACACCAATGGAATCTTGTGCTGTGTGACAGAAATGAGGCTGAG 780
Db      721 GTAGGCTTTAATGACACCAATGGAATCTTGTGCTGTGTGACAGAAATGAGGCTGAG 780
Qy      781 ATTTCTCTCTCCACCTTGATGCAATGCTGTGTGATGGAATGTTCTTTCAATTAAT 840
Db      781 ATTTCTCTCTCCACCTTGATGCAATGCTGTGTGATGGAATGTTCTTTCAATTAAT 840
Qy      841 GGAAGCCCTCAGAACACCTTTTGTACCAAGGCTGATTTGATCTTCCATCGCAGAACT 900
Db      841 GGAAGCCCTCAGAACACCTTTTGTACCAAGGCTGATTTGATCTTCCATCGCAGAACT 900
Qy      901 TTGATTTGGTGAATGATCTTCAAGAGTGTGACACCAAAATGAATCTGTGTGTTGAT 960
Db      901 TTGATTTGGTGAATGATCTTCAAGAGTGTGACACCAAAATGAATCTGTGTGTTGAT 960
Qy      961 TTTCTTTGGGGGTGATGATCAAGCAACATCTATGTTAGTTTAAACCATCTGGGAAC 1020
Db      961 TTTCTTTGGGGGTGATGATCAAGCAACATCTATGTTAGTTTAAACCATCTGGGAAC 1020
Qy      1021 AATGATGGTATGATCTCTCGGCTCCCAAAACCTTCCGCTCCAGAGAAATCTCCAAGAC 1080
Db      1021 AATGATGGTATGATCTCTCGGCTCCCAAAACCTTCCGCTCCAGAGAAATCTCCAAGAC 1080
Qy      1081 AACGTTTGTGAAGATGATGTCACACGACATGCAATCTCTTAAGAGCTGTGTGAACATCCC 1140
Db      1081 AACGTTTGTGAAGATGATGTCACACGACATGCAATCTCTTAAGAGCTGTGTGAACATCCC 1140

```

```

Qy      1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAAGGAGATACCAAGAGCCATGGAT 1200
Db      1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAAGGAGATACCAAGAGCCATGGAT 1200
Qy      1201 GAGTACATTTCAAGATATTTATGAGGTGGAAGAAACACCAATTTGTTTGCACAAACATGT 1260
Db      1201 GAGTACATTTCAAGATATTTATGAGGTGGAAGAAACACCAATTTGTTTGCACAAACATGT 1260
Qy      1261 GAGGATTCCTTTTAGCTGCTCTTATTTATCTTGGACTTGTGCTTCTTGTGAGCTGAGC 1320
Db      1261 GAGGATTCCTTTTAGCTGCTCTTATTTATCTTGGACTTGTGCTTCTTGTGAGCTGAGC 1320
Qy      1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGAAATTCACATATTCACCCAGTTGCT 1380
Db      1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGAAATTCACATATTCACCCAGTTGCT 1380
Qy      1381 ACCATTCTCAGCTATCTACCAAGGCTCTCTGTTCCACCGGGTACCAAGTGTGAT 1440
Db      1381 ACCATTCTCAGCTATCTACCAAGGCTCTCTGTTCCACCGGGTACCAAGTGTGAT 1440
Qy      1441 GCATTGTCAAGACCGCTGCAATGCTGGAAGAAACATTAATGAGGCTTGTGATTGAGCC 1500
Db      1441 GCATTGTCAAGACCGCTGCAATGCTGGAAGAAACATTAATGAGGCTTGTGATTGAGCC 1500
Qy      1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533
Db      1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533

RESULT 5
ADS82005
ID ADS82005 standard; cDNA; 1533 BP.
XX
XX ADS82005;
AC 18-NOV-2004 (first entry)
XX
XX Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 2.
DE Soybean; plant; myo-inositol 1-phosphate synthase; ssi; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX Glycine max; cultivar Wye.
OS
XX
XX Key Location/Qualifiers
FH CDS 1..1533
FT /*tag=a
FT /product="myo-inositol 1-phosphate synthase"
FT
XX
XX US2003074685-A1.
PN
XX
XX 17-APR-2003.
PD
XX
XX 11-MAR-2002; 2002US-00025003.
PF
XX
XX 08-APR-1997; 97US-00835751.
PR
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
PI
XX
XX MPI: 2004-63957/62.
DR P-PSDB; ADS82006.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 2; SEQ ID NO 15; 34bp; English.
XX

```

The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype with (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.

Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 96.7%; Score 1482; DB 13; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGTTGATGAGAAATTTAAAGTAGAGAGTCTTAATGTGAATGACCCAGATGAGATT 60
1 ATGTTGATGAGAAATTTAAAGTAGAGAGTCTTAATGTGAATGACCCAGATGAGATT 60
61 CAGTCCGTGACAACTACGAAACCAACGAACTGTTCAAGAGAAAGAAATGACCTAT 120
61 CAGTCCGTGACAACTACGAAACCAACGAACTGTTCAAGAGAAAGAAATGACCTAT 120
121 CAGTGGATTGTCAAAACCAATCCGTCACTACCAATTTAAACCAACCAATGTTTCA 180
121 CAGTGGATTGTCAAAACCAATCCGTCACTACCAATTTAAACCAACCAATGTTTCA 180
181 AATTTGGGGGTGATGCTTGTGGGTGGGAGAAACAACGCGCTTACCTTCAACGGTGGT 240
181 AATTTGGGGGTGATGCTTGTGGGTGGGAGAAACAACGCGCTTACCTTCAACGGTGGT 240
241 GTTATGCTTAACAGAGAGCATTTGATGGGCTACAAAGAGCAAGATTCAACAGCAAT 300
241 GTTATGCTTAACAGAGAGCATTTGATGGGCTACAAAGAGCAAGATTCAACAGCAAT 300
301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATCTTTCAGGAGAGAA 360
301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATCTTTCAGGAGAGAA 360
361 ACTATAGCCCAATTAAGAGTCTGCTTCAATGCTTAATCTTCAAGCAATGTTGGG 420
361 ACTATAGCCCAATTAAGAGTCTGCTTCAATGCTTAATCTTCAAGCAATGTTGGG 420
421 GGATGGATATACAGCAATGAACTGCTGATGTCATGGCCAGGCAAGAGTGTGAC 480
421 GGATGGATATACAGCAATGAACTGCTGATGTCATGGCCAGGCAAGAGTGTGAC 480
481 ATCGATTGAGAGAGAGTGAAGCTTACATGAATTCATGTTTCACTCCCGGAATC 540
481 ATCGATTGAGAGAGAGTGAAGCTTACATGAATTCATGTTTCACTCCCGGAATC 540

541 TACAGCCGGATTTTATGCTGCAACCAAGAGAGAGTCCCAACAGATGATTAGGGC 600
541 TACAGCCGGATTTTATGCTGCTCAACCAAGAGAGAGTCCCAACAGATGATTAGGGC 600
601 ACAAGCAAGAGCAAGTTCAGCAAAATCATCAAGATCAAGGCTTTAAGAAAGCACC 660
601 ACAAGCAAGAGCAAGTTCAGCAAAATCATCAAGATCAAGGCTTTAAGAAAGCACC 660
661 AAGTGAACAGAGTGTGCTTCTGTGAGTCCCAACAGAGAGTATGCAATTTGGTT 720
661 AAGTGAACAGAGTGTGCTTCTGTGAGTCCCAACAGAGAGTATGCAATTTGGTT 720
721 GTAGGCTTAATGACCAACAGAGATCTTGGGCTGTGAGCAAGAAATGAGGCTGAG 780
721 GTAGGCTTAATGACCAACAGAGATCTTGGGCTGTGAGCAAGAAATGAGGCTGAG 780
781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGATGAGAAATGTTCTTTTCAATTAAT 840
781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGATGAGAAATGTTCTTTTCAATTAAT 840
841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGCTTCCAGAGAAACT 900
841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGCTTCCAGAGAAACT 900
901 TTGATTGAGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATGTTGTTGAT 960
901 TTGATTGAGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATGTTGTTGAT 960
961 TTTCTGTGGGGCTGGTATCAAGCCCAATCATATAGTTAGTTACCAACATCTGGAAAC 1020
961 TTTCTGTGGGGCTGGTATCAAGCCCAATCATATAGTTAGTTACCAACATCTGGAAAC 1020
1021 AATGATGATATATCTTGGCTCCCAACACCTTCCGTCAGAGAAATCTTCAAGAGC 1080
1021 AATGATGATATATCTTGGCTCCCAACACCTTCCGTCAGAGAAATCTTCAAGAGC 1080
1081 AAGTTTGAAGATATGCTCAACAGCAATGCAATCTTATGAGCTGTGTAACATCCC 1140
1081 AAGTTTGAAGATATGCTCAACAGCAATGCAATCTTATGAGCTGTGTAACATCCC 1140
1141 GACATGTTGTTATTAATGATGTCCTTACGTAGGGATAGCAAGAGCAGATGAT 1200
1141 GACATGTTGTTATTAATGATGTCCTTACGTAGGGATAGCAAGAGCAGATGAT 1200
1201 GAGTACACTTCAAGATATTCATGAGTGAAGAAACCACTTGTGTTGCAACACATGT 1260
1201 GAGTACACTTCAAGATATTCATGAGTGAAGAAACCACTTGTGTTGCAACACATGT 1260
1261 GAGGATTCCTTTTACGCTCTCTATATCTTGAATGCTGCTTCTTGTGAGCTGAGC 1320
1261 GAGGATTCCTTTTACGCTCTCTATATCTTGAATGCTGCTTCTTGTGAGCTGAGC 1320
1321 ACTAGATCAGATTTAAAGCTGAAATGAGGAGAAATTCACATCTTCCACCAAGTGTCT 1380
1321 ACTAGATCAGATTTAAAGCTGAAATGAGGAGAAATTCACATCTTCCACCAAGTGTCT 1380
1381 ACCATTTCAAGTATTCAGCAAGGCTCTGTGTTTCCACCGGATACCAAGTGTGAAT 1440
1381 ACCATTTCAAGTATTCAGCAAGGCTCTGTGTTTCCACCGGATACCAAGTGTGAAT 1440
1441 GGAATGTCAAAACAGCGTCAATGCTGAAACCATTAAGAGGCTTGTGATGTTGGC 1500
1441 GGAATGTCAAAACAGCGTCAATGCTGAAACCATTAAGAGGCTTGTGATGTTGGC 1500
1501 CCAGAGAAATACATGATTTCTCAGTACCAAGTGA 1533
1501 CCAGAGAAATACATGATTTCTCAGTACCAAGTGA 1533

RESULT 6
AD582003
ID AD582003 standard; cDNA; 1533 BP.
XX

AC	ADBS82003;
AD	18-NOV-2004 (first entry)
DT	
XX	Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 4.
XX	
DE	
XX	Soybean, plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
KM	raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX	
OS	Glycine max; line 29018JUP03.
XX	
FH	Key Location/Qualifiers
FT	FT 1.1533
FT	CDS /tag= a
FT	/product= "myo-inositol 1-phosphate synthase"
PN	US2003074685-A1.
XX	
PD	17-APR-2003.
XX	
XX	11-MAR-2002; 2002US-00025003.
PF	
XX	08-APR-1997; 97US-00835751.
XX	PR 07-APR-1998; 98WO-US006822.
XX	
PA	(HITZ/) HITZ W D.
PA	(SEBA/) SEBASTIAN S A.
PI	Hitz WD, Sebastian SA;
XX	
DR	WPI: 2004-639957/62.
XX	P-PSDB; ADBS82004.
PT	
PT	Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT	phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX	phytic acid and inorganic phosphate content of soybean seeds.
XX	
XX	Example 8; SEQ ID NO 13; 34pp; English.
PS	
XX	The invention relates to an isolated nucleic acid fragment encoding a
CC	soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC	phosphate synthase having decreasing capacity for the synthesis for myo-
CC	inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC	nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC	complement, subfragment or the complement of the subfragment, operably
CC	linked to suitable regulatory sequences, where expression of the chimeric
CC	gene results in a decrease in expression of an endogenous or native gene
CC	encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC	comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC	acid content of less than 17 micromol/g), a seed content of raffinose plus
CC	stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC	greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC	from the plant, making a soybean plant with the heritable phenotype
CC	(comprising crossing LR33 or the plant comprising the chimeric gene with
CC	an elite soybean plant and selecting a progeny plant of the cross or
CC	crossing step that has a heritable phenotype as mentioned above), seeds
CC	of soybean plant made by the above method, a soy protein product derived
CC	from seeds of a soybean plant (homologous for one or more gene encoding a
CC	mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC	the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC	heritable phenotype as mentioned above), and making or producing a
CC	soybean protein product derived from seeds of a soybean plant with
CC	heritable phenotype as mentioned above. The nucleic acid is useful for
CC	altering raffinose saccharide, sucrose, phytic acid and inorganic
CC	phosphate content of soybean seeds thus leading to valuable and useful
CC	soybean products, since the presence of high concentration of raffinose
CC	oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC	when consumed by humans. The present sequence encodes a wild-type myo-
XX	inositol 1-phosphate synthase.
XX	
XX	Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 96.7%; Score 1482; DB 13; Length 1533;

Best Local Similarity 99.9%: Pred No. 0; Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	1	ATGTTTCATCGAAGATTTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGACTGAGATT	60						
Db	1	ATGTTTCATCGAAGATTTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGACTGAGATT	60						
QY	61	CAGTCCGCTGATCACTACGAAACACCGAACTTGTTACGAGAACAGAAATGGCACTAT	120						
Db	61	CAGTCCGCTGATCACTACGAAACACCGAACTTGTTACGAGAACAGAAATGGCACTAT	120						
QY	121	CAGTGAATGTCGAACCAACCAATCCGCTACACTACCAATTTAAACCAACACCACTGTTCCA	180						
Db	121	CAGTGAATGTCGAACCAACCAATCCGCTACACTACCAATTTAAACCAACCACTGTTCCA	180						
QY	181	AAATTTGGGGGTGATGCTTGTGGGGTTGGGGTGGAAAACAACGGCTCTACCGCTGAGT	240						
Db	181	AAATTTGGGGGTGATGCTTGTGGGGTTGGGGTGGAAAACAACGGCTCTACCGCTGAGT	240						
QY	241	GTTATTTGCTAACAGAGAGACATTTTCATGGGGCTTCAAGAGACAAAGATTTCAACAGCAAT	300						
Db	241	GTTATTTGCTAACAGAGAGAGATTTTCATGGGGCTTCAAGAGACAAAGATTTCAACAGCAAT	300						
QY	301	TACTTTGGGCTCCCTCACCAAGCCTCAGCTATTGAGTTGGATCTCTTCAAGGAGAGAAA	360						
Db	301	TACTTTGGGCTCCCTCACCAAGCCTCAGCTATTGAGTTGGATCTCTTCAAGGAGAGAAA	360						
QY	361	ATCTATGCCCCATTCAGAGAGTCTGCTTCCAAATGTTAATCCTGACGACATTTGTGTTGGG	420						
Db	361	ATCTATGCCCCATTCAGAGAGTCTGCTTCCAAATGTTAATCCTGACGACATTTGTGTTGGG	420						
QY	421	GGATGGGATATCAGGACATGAACCTGCGTATGCCATGGCCAGGGCAAGGTGTTGAC	480						
Db	421	GGATGGGATATCAGGACATGAACCTGCGTATGCCATGGCCAGGGCAAGGTGTTGAC	480						
QY	481	ATCGATTTGCGAAGACGTTGAGGGCTTTACATGGAATCCATGTTCCATGCCCGCGAATC	540						
Db	481	ATCGATTTGCGAAGACGTTGAGGGCTTTACATGGAATCCATGTTCCATGCCCGCGAATC	540						
QY	541	TACGACCCGGATTTTCATTGCTGCCCAACCAAGAGAGCGTGCACAAACGTGATTAAAGGC	600						
Db	541	TACGACCCGGATTTTCATTGCTGCCCAACCAAGAGAGCGTGCACAAACGTGATTAAAGGC	600						
QY	601	ACAAAGCAGAGCAGATTACGCAAAATCATCAAGACATCAAGGCGTTTAAAGAACCAAC	660						
Db	601	ACAAAGCAGAGCAGATTACGCAAAATCATCAAGACATCAAGGCGTTTAAAGAACCAAC	660						
QY	661	AAAGTGCACAAAGGTGTTGCTCTGTGACATGCCCAACAGAGAGTAPAGCAATTTGGTT	720						
Db	661	AAAGTGCACAAAGGTGTTGCTCTGTGACATGCCCAACAGAGAGTAPAGCAATTTGGTT	720						
QY	721	GTAGGCTTATGACACCATGAGAAATCTCTGGCTGCTGTGACACAGAAATGAGGCTGAG	780						
Db	721	GTAGGCTTATGACACCATGAGAAATCTCTGGCTGCTGTGACACAGAAATGAGGCTGAG	780						
QY	781	ATTTCTCTCTTCACCTTGTATGCCATTGCTGTGTGATGAAAAATGTTCTTTTCATTAAAT	840						
Db	781	ATTTCTCTCTTCACCTTGTATGCCATTGCTGTGTGATGAAAAATGTTCTTTTCATTAAAT	840						
QY	841	GGAGCCCTCGAACACTTTTGTACAGGGCTGATTAATCTTGCAATCGCAGGAAACACT	900						
Db	841	GGAGCCCTCGAACACTTTTGTACAGGGCTGATTAATCTTGCAATCGCAGGAAACACT	900						
QY	901	TTGATTTGGTGAATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTTGAT	960						
Db	901	TTGATTTGGTGAATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTTGAT	960						
QY	961	TTTCTTTGGGGGGCTGTGATCAAGCCCAACATCTATAGTTACAAACATCTGGGAAC	1020						
Db	961	TTTCTTTGGGGGGCTGTGATCAAGCCCAACATCTATAGTTACAAACATCTGGGAAC	1020						
QY	1021	AATGATGATGAATCTCTGGGCTTCCAAACCTTCGGCTTCAAGAAATCTCCAAAGC	1080						

DB 1021 AATGATGATGATGATCTCTGGCTCCACAAACCTTCGCTCCAAAGAAATCTCCAAAGAGC 1080
 QY 1081 AAGTGTGTTGACATATGCTCAACAGAAATGCTCTATGAGCTGTGTGAACATCCC 1140
 DB 1081 AAGTGTGTTGACATATGCTCAACAGAAATGCTCTATGAGCTGTGTGAACATCCC 1140
 QY 1141 GACCATGTTGTTTATTAAGTATGCTCTTACGTAAGGAGATGCAAGAGCATGGAT 1200
 DB 1141 GACCATGTTGTTTATTAAGTATGCTCTTACGTAAGGAGATGCAAGAGCATGGAT 1200
 QY 1201 GAGTACACTTCAAGATATTCATGGGTGAAGAAACATCTGTTTGAACACATGT 1260
 DB 1201 GAGTACACTTCAAGATATTCATGGGTGAAGAAACATCTGTTTGAACACATGT 1260
 QY 1261 GAGGATTCCTTTAGTGTCTCTATTTATCTTGAAGCTGGTCTTGTGCTGAGCTGAGC 1320
 DB 1261 GAGGATTCCTTTAGTGTCTCTATTTATCTTGAAGCTGGTCTTGTGCTGAGCTGAGC 1320
 QY 1321 ACTAGATCCAGTTTAAAGCTGAAGAAATGAGGAAATTCGACTCATTCACCGAGTTGCT 1380
 DB 1321 ACTAGATCCAGTTTAAAGCTGAAGAAATGAGGAAATTCGACTCATTCACCGAGTTGCT 1380
 QY 1381 ACCATTCCTAGCTATCTGACCAAGGCTCTCTGTTCCACCGGATCACCAAGTGTGAAT 1440
 DB 1381 ACCATTCCTAGCTATCTGACCAAGGCTCTCTGTTCCACCGGATCACCAAGTGTGAAT 1440
 QY 1441 GCATTTGCAAGAGCGGTGCAATGCTGGAAGAAACATATGAGGCTGTGTGATTGGCC 1500
 DB 1441 GCATTTGCAAGAGCGGTGCAATGCTGGAAGAAACATATGAGGCTGTGTGATTGGCC 1500
 QY 1501 CCAGAGATTAACATGATTTCTGAGTCAAGTGA 1533
 DB 1501 CCAGAGATTAACATGATTTCTGAGTCAAGTGA 1533

RESULT 7
 AAV62443
 ID AAV62443 standard; cDNA; 1533 BP.

AC AAV62443;
 XX 17-OCT-2003 (revised)
 DT 02-FEB-1999 (first entry)
 XX
 DE Soybean mutant myo-inositol 1-phosphate synthase cDNA.
 XX
 KM Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
 KW phytic acid; dg.
 XX
 OS Glycine max; line LR33.
 XX
 PN WO9845448-A1.
 XX
 PD 15-OCT-1998.
 XX
 PF 07-APR-1998; 98WO-US006822.
 XX
 PR 08-APR-1997; 97US-00835751.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Hitz WD, Sebastian SA;
 XX
 DR WPI; 1998-568353/48.
 DR P-PSDB; AAW79741.
 XX
 PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
 PT for generating plants with altered levels of e.g. raffinose, stachyose,
 PT phytic acid, etc.
 XX
 PS Example 5; Page 48-49; 63pp; English.
 CC This is the nucleotide sequence of cDNA encoding a mutant soybean myo-

CC inositol 1-phosphate synthase (MI 1-PS) present in clone LR33-10. This
 CC clone was isolated from a cDNA library of soybean line LR33 by PCR
 CC amplification (see AAV62441-42). Line LR33 was obtained by chemical
 CC mutagenesis of wild-type soybean genome and as a reduced raffinose
 CC saccharide phenotype. MI 1-PS is involved in glucose metabolism to phytic
 CC acid, raffinose and stachyose. Sequencing revealed a single base change
 CC mutation (G to T at base 1241) in the LR33 sequence when compared to the
 CC wild-type sequence (see AAV62440). The mutation results in a seed
 CC phenotype of very low raffinose saccharide sugars, very high sucrose and
 CC low phytic acid. The mutated MI 1-PS nucleic acid is used to alter the
 CC raffinose saccharide, sucrose, phytic acid and inorganic phosphate
 CC content of soybean seeds, leading to useful soybean products, e.g. a seed
 CC phytic acid content of less than 17 ug/g, a seed content of raffinose and
 CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
 CC greater than 200 ug/g. (Updated on 17-Oct-2003 to standardise OS field)
 XX
 SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 9.5%; Score 146; DB 2; Length 1533;
 Best Local Similarity 100.0%; Pred. No. 1e-60; Mismatches 0; Gaps 0;
 Matches 146; Conservative 0; Indels 0;

QY 817 ATGAAATGTTCTTTCAATTAATGAAAGCCCTCAGAACATTTGTACAGAGCTGAT 876
 DB 817 ATGAAATGTTCTTTCAATTAATGAAAGCCCTCAGAACATTTGTACAGAGCTGAT 876
 QY 877 GATCTTGCATGCGAGAACACTTTGATGATGATGACTTCAAGAGTGTCAAGC 936
 DB 877 GATCTTGCATGCGAGAACACTTTGATGATGATGACTTCAAGAGTGTCAAGC 936
 QY 937 AAAATGAATCTGTGTGTTGATT 962
 DB 937 AAAATGAATCTGTGTGTTGATT 962

RESULT 8
 ADQ14498
 ID ADQ14498 standard; cDNA; 1533 BP.

AC ADQ14498;
 XX 23-SEP-2004 (first entry)
 DT
 XX
 DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #2.
 XX
 KM Soybean; myo-inositol 1-phosphate synthase; gene; ss;
 KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
 XX inorganic phosphate; mutant.
 XX
 OS Glycine max.
 XX
 PN Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1533
 FT /*tag= a
 FT /product= "Mutant soybean myo-inositol 1-phosphate
 FT synthase #2"
 XX
 PN US2004128713-A1.
 XX
 PD 01-JUL-2004.
 XX
 PF 21-NOV-2003; 2003US-00718952.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 PR 26-APR-1999; 99US-00299315.
 PR 11-MAR-2002; 2002US-00025003.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PA (GRAC/) GRACE D J.
 PA (STREIT/) STREIT L G.


```
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
PI WPI: 2004-53135/51.
XX P-PSDB; ADQ14499.
DR
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX Example 8; SEQ ID NO 9; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;
Query Match 9.5%; Score 146; DB 12; Length 1533;
Best Local Similarity 100.0%; Pred. NO. 1e-60; Mismatches 0; Gaps 0;
Matches 146; Conservative 0; Indels 0;
QY 817 ATGGAATAATGTTCTTTCATTATATGAGAGCCCTCAGAACCTTTGTACAGGGCTGATT 876
DB 817 ATGGAATAATGTTCTTTCATTATATGAGAGCCCTCAGAACCTTTGTACAGGGCTGATT 876
QY 877 GATCTTGCATCGCGAGAGAACCTTTGATTGGTGAGATGACTTCAAGAGTGTCAGACC 936
DB 877 GATCTTGCATCGCGAGAGAACCTTTGATTGGTGAGATGACTTCAAGAGTGTCAGACC 936
QY 937 AAAATGAATCTGTGTGGTTGATT 962
DB 937 AAAATGAATCTGTGTGGTTGATT 962
RESULT 9
ADQ14494
ID ADQ14494 standard; cDNA, 1533 BP.
XX
XX ADQ14494;
AC 23-SEP-2004 (first entry)
DT
XX
XX Mutant soybean myo-inositol 1-phosphate synthase cDNA #1.
XX
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate; mutant.
OS Glycine max.
XX Synthetic.
XX
XX Key Location/Qualifiers
FH 1..1533
FT CDS /*tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT /synthase #1"
```

```
FT mutation replace(1241,G)
XX /*tag= b
XX
XX US2004128713-A1.
XX
XX 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98MO-US006882.
XX 26-APR-1999; 99US-00289315.
XX 11-MAR-2002; 2002US-00025003.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX (GRAC/) GRACE D J.
XX (STRE/) STREIT L G.
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
PI WPI: 2004-53135/51.
XX P-PSDB; ADQ14495.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX Claim 10; SEQ ID NO 5; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;
Query Match 9.5%; Score 146; DB 12; Length 1533;
Best Local Similarity 100.0%; Pred. NO. 1e-60; Mismatches 0; Gaps 0;
Matches 146; Conservative 0; Indels 0;
QY 817 ATGGAATAATGTTCTTTCATTATATGAGAGCCCTCAGAACCTTTGTACAGGGCTGATT 876
DB 817 ATGGAATAATGTTCTTTCATTATATGAGAGCCCTCAGAACCTTTGTACAGGGCTGATT 876
QY 877 GATCTTGCATCGCGAGAGAACCTTTGATTGGTGAGATGACTTCAAGAGTGTCAGACC 936
DB 877 GATCTTGCATCGCGAGAGAACCTTTGATTGGTGAGATGACTTCAAGAGTGTCAGACC 936
QY 937 AAAATGAATCTGTGTGGTTGATT 962
DB 937 AAAATGAATCTGTGTGGTTGATT 962
RESULT 10
ADS81999
ID ADS81999 standard; cDNA, 1533 BP.
XX
```


Thu Jun 9 09:20:47 2005

AC ADS81999;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 3.
XX
KM Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; line 29004UP01.
XX
XX
XX Key Location/Qualifiers
FH 1..1533
FT /*tag= a
FT /product= "myo-inositol 1-phosphate synthase"
XX
XX US2003074685-A1.
XX
PD 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX MPI: 2004-639957/62.
XX P-PSDB; ADS82000.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Example 8; SEQ ID NO 9; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
XX soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX phosphate synthase having decreasing capacity for the synthesis for myo-
XX inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX complement, subfragment or the complement of the subfragment, operably
XX linked to suitable regulatory sequences, where expression of the chimeric
XX gene results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX comprising the chimeric gene (with a heritable phenotype of a seed phytic
XX acid content of less than 17 micromol/g, a seed content of raffinose plus
XX stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX greater than 200 micromol/g, provided that the plant is not LR33), seeds
XX from the plant, making a soybean plant with the heritable phenotype
XX (comprising crossing LR33 or the plant comprising the chimeric gene with
XX an elite soybean plant and selecting a progeny plant of the cross of
XX crossing step that has a heritable phenotype as mentioned above), seeds
XX of soybean plant made by the above method, a soy protein product derived
XX from seeds of a soybean plant (homologous for one or more gene encoding a
XX mutant myo-inositol 1-phosphate synthase having decreased capacity for
XX the synthesis of myo-inositol 1-phosphate, where the gene confers a
XX heritable phenotype as mentioned above), and making or producing a
XX soybean protein product derived from seeds of a soybean plant with a
XX heritable phenotype as mentioned above. The nucleic acid is useful for
XX altering raffinose saccharide, sucrose, phytic acid and inorganic
XX phosphate content of soybean seeds thus leading to valuable and useful
XX soybean products, since the presence of high concentration of raffinose
XX oligosaccharides in soy plants (and other legumes) can lead to flatulence
XX when consumed by humans. The present sequence encodes a wild-type myo-
XX inositol 1-phosphate synthase.
XX
XX Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;
XX
XX Query Match 9.5%; Score 146; DB 13; Length 1533;

Best Local Similarity 100.0%; Pred. No. 1e-60;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 817 ATGAAAATGCTCTTCAATTAATGAGCCCTCAGACACTTTGTACAGGCTGATT 876
DB 817 ATGAAAATGCTCTTCAATTAATGAGCCCTCAGAACACTTTGTACAGGCTGATT 876
OY 877 GATCTTGCATGCGCAGGAACACTTTGATTGAGATGACTTCAAGTGTCTGACAC 936
DB 877 GATCTTGCATGCGCAGGAACACTTTGATTGAGATGACTTCAAGTGTCTGACAC 936
OY 937 AAAATGAATCTGTGTGTTGATT 962
DB 937 AAAATGAATCTGTGTGTTGATT 962
RESULT 11
ADS81993
ID ADS81993 standard; cDNA; 1533 BP.
XX
XX ADS81993;
XX
XX 18-NOV-2004 (first entry)
XX
XX
XX Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 1.
XX
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
XX Glycine max; cultivar Wye.
XX
XX
XX Key Location/Qualifiers
FH 1..1533
FT /*tag= a
FT /product= "myo-inositol 1-phosphate synthase"
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX MPI: 2004-639957/62.
XX P-PSDB; ADS81994.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 2; SEQ ID NO 1; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
XX soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX phosphate synthase having decreasing capacity for the synthesis for myo-
XX inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX complement, subfragment or the complement of the subfragment, operably
XX linked to suitable regulatory sequences, where expression of the chimeric
XX gene results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX comprising the chimeric gene (with a heritable phenotype of a seed phytic
XX acid content of less than 17 micromol/g, a seed content of raffinose plus
XX stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX greater than 200 micromol/g, provided that the plant is not LR33), seeds
XX from the plant, making a soybean plant with the heritable phenotype

OS Glycine max.
 XX
 FH Key Location/Qualifiers
 FT CDS 54..1586
 FT /tag= a
 FT /product= "Wild type soybean myo-inositol 1-phosphate
 FT synthase #1"
 XX
 XX US2004128713-A1.
 XX
 XX
 XX
 XX
 XX 01-JUL-2004.
 XX
 XX 21-NOV-2003; 2003US-00718952.
 XX
 XX
 XX 08-APR-1997; 97US-00835751.
 XX 07-APR-1998; 98WO-US006822.
 XX 26-APR-1999; 99US-00299315.
 XX 11-MAR-2002; 2002US-00025003.
 XX
 XX (HITZ/) HITZ W D.
 XX (SEBA/) SEBASTIAN S A.
 XX (GRAC/) GRACE D J.
 XX (STRE/) STREIT L G.
 XX
 XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
 XX
 XX WPI; 2004-533135/51.
 XX P-PSDB; ADQ14491.
 XX
 XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 XX useful for producing plants with decreased raffinose, stachyose, and
 XX phytic acid and increased sucrose, leading to valuable and useful soybean
 XX products.
 XX
 XX Claim 4; SEQ ID NO 1; 48pp; English.
 XX
 XX
 XX The invention relates to a nucleic acid fragment encoding a soybean myo-
 XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
 XX phosphate. The invention also relates to a chimeric gene operably linked
 XX to suitable regulatory sequences, where expression of the chimeric gene
 XX results in a decrease in expression of an endogenous or native gene
 XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
 XX comprising the chimeric gene, a method of making the soybean plant, a
 XX seed of the soybean plant, a soy protein product derived from the
 XX processing of soybean seeds, a method of making or producing a soy
 XX protein product and a method of using a soybean plant homozygous for at
 XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 XX nucleic acid fragment and methods are useful for producing plants with
 XX decreased raffinose, stachyose and phytic acid content and increased
 XX sucrose and inorganic phosphate content, leading to valuable and useful
 XX soybean products. This sequence represents cDNA encoding a wild type
 XX soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
 XX
 XX Sequence 1760 BP; 494 A; 371 C; 411 G; 484 T; 0 U; 0 Other;
 XX
 XX Query Match 9.5%; Score 146; DB 12; Length 1760;
 XX Best Local Similarity 100.0%; Pred. No. 1e-60;
 XX Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 817 ATGGAATAATGTTCTTCAATTAATGAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
 XX |||||||
 XX DB 870 ATGGAATAATGTTCTTCAATTAATGAGCCCTCAGAACACTTTTGTACAGGCGTGATT 929
 XX |||||||
 XX QY 877 GATCTTGCCATCGCGAGGAACAATTGATTGGTGAGATGATCTTCAAGAGTGTGACACC 936
 XX |||||||
 XX DB 930 GATCTTGCCATCGCGAGGAACAATTGATTGGTGAGATGATCTTCAAGAGTGTGACACC 989
 XX |||||||
 XX QY 937 AAAATGAATCTGTGTGTTGATT 962
 XX |||||||
 XX DB 990 AAAATGAATCTGTGTGTTGATT 1015
 XX |||||||

RESULT 14
 AAV62440.
 ID AAV62440 standard; cDNA; 1782 BP.
 XX
 XX AAV62440;
 XX
 XX
 XX 17-OCT-2003 (revised)
 XX 02-FEB-1999 (first entry)
 XX
 XX Soybean wild-type myo-inositol 1-phosphate synthase cDNA.
 XX
 XX
 XX Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
 XX phytic acid; ds.
 XX
 XX Glycine max; line LR13.
 XX
 XX
 XX
 XX
 XX W09845448-A1.
 XX
 XX 15-OCT-1998.
 XX
 XX 07-APR-1998; 98WO-US006822.
 XX
 XX 08-APR-1997; 97US-00835751.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 XX Hitz WD, Sebastian SA;
 XX
 XX WPI; 1998-568353/48.
 XX P-PSDB; AAW79740.
 XX
 XX Soybean plants containing altered myo-inositol-1-phosphate gene - useful
 XX for generating plants with altered levels of e.g. raffinose, stachyose,
 XX phytic acid, etc.
 XX
 XX Example 5; Page 44-45; 63pp; English.
 XX
 XX This is the nucleotide sequence of cDNA encoding the wild-type soybean
 XX myo-inositol 1-phosphate synthase (MI 1-PS) present in clone p5pm1-1ps
 XX (ATCC 97970). The clone was isolated from a cDNA library of soybean line
 XX LR13 by hybridisation to a probe made from MI 1-PS cDNA of Arabidopsis
 XX thaliana. MI 1-PS is involved in glucose metabolism to phytic acid,
 XX raffinose and stachyose. A mutant MI 1-PS nucleic acid (see AAV62443) has
 XX been identified in soybean line LR13, a mutagenised line of low raffinose
 XX saccharide phenotype. Sequencing revealed a single base change mutation
 XX (G to T at base 1241) in the LR13 sequence. The mutation results in a
 XX seed phenotype of very low raffinose saccharide sugars, very high sucrose
 XX and low phytic acid. The nucleic acid is used to alter the raffinose
 XX saccharide, sucrose, phytic acid and inorganic phosphate content of
 XX soybean seeds, leading to useful soybean products, e.g. a seed phytic
 XX acid content of less than 17 ug/g, a seed content of raffinose and
 XX stachyose combined of less than 14.5 ug/g, and a seed sucrose content
 XX greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 XX Sequence 1782 BP; 516 A; 371 C; 411 G; 484 T; 0 U; 0 Other;
 XX
 XX Query Match 9.5%; Score 146; DB 2; Length 1782;
 XX Best Local Similarity 100.0%; Pred. No. 1e-60;
 XX Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 817 ATGGAATAATGTTCTTCAATTAATGAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
 XX |||||||
 XX DB 870 ATGGAATAATGTTCTTCAATTAATGAGCCCTCAGAACACTTTTGTACAGGCGTGATT 929
 XX |||||||
 XX QY 877 GATCTTGCCATCGCGAGGAACAATTGATTGGTGAGATGATCTTCAAGAGTGTGACACC 936
 XX |||||||
 XX DB 930 GATCTTGCCATCGCGAGGAACAATTGATTGGTGAGATGATCTTCAAGAGTGTGACACC 989
 XX |||||||

QY 937 AAAATGAATCTGTGTGTTGATT 962
 DB 990 AAAATGAATCTGTGTGTTGATT 1015

DB 95 ATCCTATAGAGCTGTGAACATCC 120
 Search completed: June 8, 2005, 07:27:00
 Job time : 843.149 secs

RESULT 15

ID ADP93370 standard; cDNA; 377 BP.
 AC ADP93370;

DT 09-SEP-2004 (first entry)

DE Cotton expressed sequence tag, EST, #2381.

KM Cotton; ser; EST; expressed sequence tag; plant; plant protection;
 KM plant improvement; marker-assisted breeding.

OS Gossypium hirsutum; variety Nucleon33B.

XX US2004123338-A1.

PN 24-JUN-2004.

PF 08-DEC-2000; 2000US-00732627.

PR 10-DEC-1999; 99US-0170255P.

XX (FINC/) FINCHER K L.

PA Finch K L.

PI Finch K L.

XX WPI; 2004-479807/45.

PT New substantially purified nucleic acid molecule that encodes a cotton
 PT protein or its fragment, useful as molecular tool for the targeting and
 PT isolation of novel genes for plant protection and improvement.

PS Claim 1, SEQ ID NO 2381; 30pp; English.

CC The invention relates to a substantially purified nucleic acid molecule
 CC that encodes a cotton protein or its fragment comprising an EST
 CC (expressed sequence tag) appearing as ADP90990-ADP95919. Also included
 CC are a substantially purified cotton protein or its fragment encoded by a
 CC nucleic acid molecule above and a transformed plant (having a nucleic
 CC acid molecule which comprises: an exogenous promoter region which
 CC functions in a plant cell to cause the production of a mRNA molecule; a
 CC structural nucleic acid molecule comprising one of the ESTs or their
 CC complements; a 3' non-translated sequence that functions in the plant
 CC cell to cause termination of transcription and addition of polyadenylated
 CC ribonucleotides to a 3' end of the mRNA molecule). The ESTs are useful as
 CC molecular tool for the targeting and isolation of novel genes for plant
 CC protection and improvement. The ESTs are useful for developing new
 CC strategies for understanding critical plant developmental and metabolic
 CC pathways, for isolating genes and promoters, for identifying and mapping
 CC the genes involved in developmental and metabolic pathways, and for
 CC determining gene function. The cotton nucleic acid molecules are useful
 CC as molecular tags to isolate genetic regions, isolate genes, map genes,
 CC and determine gene function. The nucleic acid molecules are useful for
 CC determining if genes are members of a particular gene family and for use
 CC in marker-assisted breeding programs. The present sequence is one of the
 CC 4930 cotton ESTs of the invention. NOTE: The sequences are not displayed
 CC in the specification but are available in electronic format from the
 CC USPTO at seqdata.uspto.gov/sequence.html?DocID=20040123338.

XX Sequence 377 BP; 100 A; 87 C; 90 G; 100 T; 0 U; 0 Other;

Query Match 1.7%; Score 26; DB 12; Length 377;
 Best Local Similarity 100.0%; Pred. No. 0.064;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1114 ATCCTATAGAGCTGTGAACATCC 1139
 |||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:49:46 ; Search time 5221.88 Seconds
(without alignments)
11174.613 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533
Sequence: 1 atgcttcacgcagattttaa.....tgattctcgcagacaaatga 1533

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Listing first 45 summaries

Database :

EST :
1: gb_esc1:*
2: gb_esc2:*
3: gb_hnc:*
4: gb_esc3:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	551	35.9	552	5	BM955039 sam76c03.
2	544	35.5	754	4	B1968101 GM830004A
3	543	35.4	594	4	BM523521 sam85a05.
4	540	35.2	540	5	BM887128 sam34e09.
5	517	33.7	561	1	A1442485 aa27f05.Y
6	503	32.8	525	2	BE191344 en71c11.Y
7	496	32.4	541	4	BM085458 sa337a06.
8	479	31.2	782	2	AW348136 GM210001A
9	466	30.4	585	4	BG239011 ba857d09.
10	462	30.1	618	2	AW279066 sg07d02.Y
11	444	29.0	449	2	AM203517 ef35p05.Y
12	419	27.3	632	6	CA938527 ba334a11.
13	413	26.9	516	4	BG652636 ba669g02.
14	377	24.6	505	2	BE660322 1274 Gmax
15	374	24.4	588	5	BUS49101 GM880018B
16	348	22.7	420	4	BM092617 ba16f010.
17	304	19.8	304	4	BF425513 au56f04.Y
18	297	19.4	493	5	BM081344 sam23b12.
19	282	19.0	634	2	BE331050 bc93e10.Y
20	278	18.1	457	2	AW310369 bf35p05.X
21	277	18.1	411	2	AM099866 bd17e12.Y
22	277	18.1	456	2	AM620966 ej50e08.Y
23	274	17.9	274	2	AM153049 ba34c07.Y
24	264	17.2	467	1	A1442850 ba27f05.X

25	247	16.1	670	4	BG044525 saa29e07.
26	232	15.1	539	2	AM184796 se82f07.Y
27	214	14.0	526	4	B1894109 ba16c05.Y
28	209	13.6	226	1	A1960904 bc92c07.Y
29	187	12.2	295	4	B1944254 ba95f07.Y
30	186	12.1	187	2	BF596401 bu71g11.Y
31	185	12.1	388	2	AM432728 bh85f08.Y
32	179	11.7	475	6	CB063430 ba85e09.Y
33	152	9.9	152	2	BF596037 bu68a09.Y
34	149	9.7	172	2	BE609610 bq50a08.Y
35	146	9.5	308	2	AW705757 sk51e05.Y
36	146	9.5	443	5	BQ612082 ba277b07.Y
37	146	9.5	533	4	BM523576 sam85g04.Y
38	146	9.5	661	2	BE191464 en76h10.Y
39	146	9.5	813	7	CK768601 GM-t1030.Y
40	135	8.8	290	2	AW460108 b110d02.Y
41	135	8.8	316	2	AM398011 sg71b06.Y
42	135	8.8	396	2	AW472088 b119g09.Y
43	135	8.8	619	2	BE331363 ba09f01.Y
44	135	8.8	686	2	AM348857 GM210010A
45	132	8.6	271	2	AM568795 b161d10.Y

ALIGNMENTS

RESULT 1
LOCUS BM955039 552 bp mRNA linear EST 05-JUL-2004
DEFINITION sam76c03.Y1 Gm-cl069 Glycine max cDNA clone SOYBEAN CLONE ID: GM-cl069-5118 5' similar to TR:Q9SSV4 Q9SSV4 MYO-INOSITOL-1-PHOSPHATE SYNTHASE. ; mRNA sequence.

ACCESSION BM955039
VERSION BM955039.1 GI:19453629
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE
AUTHORS 1 (bases 1 to 552)
Shoemaker,R., Keim,P., Vodkin,L., Expelding,J., Corryell,V., Khanna,A., Bolla,B., Matra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Treising,B., Allen,M., Bowers,Y., Pearson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritzer,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence etop: 420.
Location/Qualifiers

FEATURES

source
1..552
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl069-5118"
/tissue_type="Degenerating cotyledons, 9-10 day old etiolated seedling"
/lab_host="DH10B"

/clone.lib="Gm-cl069"
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI. The cDNA library was constructed from mRNA isolated from degenerating cotyledons of 9-10 day old etiolated seedlings for the cultivar Williams. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 35.5%; Score 551; DB 5; Length 552;
 Best Local Similarity 100.0%; Pred. No. 1.4e-304;
 Matches 551; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 682 CTTGTGACCTGCCAACAAGAGAGTATGACATTTGTTGAGGCTTATGACACCATG 741
DB 1 CTTGTGACCTGCCAACAAGAGAGTATGACATTTGTTGAGGCTTATGACACCATG 60
QY 742 GAGAACTCTTGGCTGCTGTGACAGAAATGAGGCTGAGATTCTCTTCCACTTGTAT 801
DB 61 GAGAACTCTTGGCTGCTGTGACAGAAATGAGGCTGAGATTCTCTTCCACTTGTAT 120
QY 802 GCCATTGCTGCTGTGATGAGAAATGTTCTTTCATTATGAGAGCCCTGACACTTTT 861
DB 121 GCCATTGCTGCTGTGATGAGAAATGTTCTTTCATTATGAGAGCCCTGACACTTTT 180
QY 862 GTACCAAGGCTGATGATCTTGCATCGGAGAGAACATTTGATGAGAGATGACTTC 921
DB 181 GTACCAAGGCTGATGATCTTGCATCGGAGAGAACATTTGATGAGAGATGACTTC 240
QY 922 AAGAGTGTGACAGCAAAATGAAATCTGTGTTGATTTCTGTGGGGCTGGTATC 981
DB 241 AAGAGTGTGACAGCAAAATGAAATCTGTGTTGATTTCTGTGGGGCTGGTATC 300
QY 982 AAGCAACATCTATAGTTAGTTACCAACATCTGGGAAACATGATGATGATCTTCG 1041
DB 301 AAGCAACATCTATAGTTAGTTACCAACATCTGGGAAACATGATGATGATCTTCG 360
QY 1042 GGTCCACAACCTCCGCTCCCAAGGAAATCTCAAGAGCAAGTTGATGATGATGTC 1101
DB 361 GGTCCACAACCTCCGCTCCCAAGGAAATCTCAAGAGCAAGTTGATGATGATGTC 420
QY 1102 AACAGCAATGCCATCTCTATAGAGCTGTGTAACATCCGACCATGTTGTTATTAG 1161
DB 421 AACAGCAATGCCATCTCTATAGAGCTGTGTAACATCCGACCATGTTGTTATTAG 480
QY 1162 TATGTGCTTATGATGAGGAGTATGCAAGAGCATGATGATGATGATGATGATTC 1221
DB 481 TATGTGCTTATGATGAGGAGTATGCAAGAGCATGATGATGATGATGATGATTC 540
QY 1222 ATGGGTGAGAA 1232
DB 541 ATGGGTGAGAA 551

```

RESULT 2
 BI968101/c 754 bp mRNA linear EST 23-OCT-2001

LOCUS BI968101 754 bp mRNA linear EST 23-OCT-2001
 DEFINITION GW830004A22D05 Gm-r1083 Glycine max cDNA clone Gm-r1083-1306 3',
 mRNA sequence.

ACCESSION BI968101

VERSION BI968101.1 GI:16342506

KEYWORDS EST.

SOURCE Glycine max (soybean)

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

roside; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 754)
 REFERENCES
 Vokkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Corvett,L.V.,
 Erpelting,J., Raph,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.
 A Functional Genomics Program for Soybean (NSF 9872565)
 Unpublished (1999)
 Other ESTs: A1855435 corresponding to Gm-cl013-777 (5')
 Contact: Vokkin, L.O., PI, A Functional Genomics Program for
 Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional
 Genomics
 University of Illinois
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l-vokkin@uiuc.edu

This clone is available through: Incyte Genomics, 4633 World
 Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or
 (314) 427-3222 FAX: (314) 427-3324. Web site:
 http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboratio
 n/index
 Seq primer: 5'-TTTCTTTTCTTTTCTTTTCTTTT(A/C/G)-3'.

FEATURES

source

Location/Qualifiers
 1..754
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone.lib="Gm-r1083-1306"
 /clone.lib="Gm-r1083"
 /note="The library Gm-r1083 is a sequence-driven, reracked set of 4,992 clones selected from cDNA libraries from various tissues and stages of development of soybean. It represents 1117 sequences from the progenitor library Gm-cl009 (from mature roots of 2 month old greenhouse grown 'Williams' soybean plants); 820 sequences from the progenitor library Gm-cl013 (from 2 to 3 week old whole plants of Williams); and 3055 sequences from library Gm-cl028 (from 'Superpod' plants whose seedlings were inoculated with Bradyrhizobium japonicum, courtesy of Dr. Gary Stacey). The 5' ESTs of the source clones from the different progenitor libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1083. The cDNA clones of the reracked Gm-r1083 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://web.auc.umn.edu/biodata/netsoy/. Reracking was performed by Incyte Genomics, St. Louis, http://www.incyte.com, and 3' sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.lie.uiuc.edu/biocech/keck.html.
 Note: The corresponding 5' EST from each clone in the Gm-r1083 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under 'OTHER EST'."

ORIGIN

Query Match 35.5%; Score 544; DB 4; Length 754;
 Best Local Similarity 100.0%; Pred. No. 1.6e-300;
 Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 990 ATCTATAGTTAGTATCAACCATCTGGGAAACATATGATGATGATCTCTGGCTCCACA 1049
DB 702 ATCTATAGTTAGTATCAACCATCTGGGAAACATATGATGATGATCTCTGGCTCCACA 643
QY 1050 AACCTTCGCTCCAGGAATCTCAAGAGCAAGCTGTGACGATATGATGATGATGATGAT 1109
DB 642 AACCTTCGCTCCAGGAATCTCAAGAGCAAGCTGTGACGATATGATGATGATGATGATGAT 583

```


Accession MYO-INSITOL-1-PHOSPHATE SYNTHASE. ;, mRNA sequence.
Version BM887128.1 GI:19270872
Keywords EST.
Source Glycine max (soybean)
Organism Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
1 (bases 1 to 540)
Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 421.
Location/Qualifiers
1..540
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl068-6738"
/tissue_type="leaf, drought stressed, 1 month old plants, greenhouse grown"
/lab_host="DH10B"
/clone_lib="Gm-cl068"
/note="Vector: pBluescript II SK+; Site 1: EcoRI, Site 2: XhoI. The cDNA library was constructed from mRNA isolated from drought stressed leaf tissue of the cultivar Williams 82. The month old greenhouse grown plants were deprived of water for 3 days prior to harvesting the stressed leaf tissue. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

Query Match 35.2%; Score 540; DB 5; Length 540;
Best Local Similarity 100.0%; Pred. No. 3e-298;
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

598 GGCACAAAGCAAGAGCAAGTTGACGCAATCATCAAGACATCAAGCGCTTTAAGAGAGCC 657
1 GGCACAAAGCAAGAGCAAGTTGACGCAATCATCAAGACATCAAGCGCTTTAAGAGAGCC 60
658 ACCAAAGTGGACAAGGTGTTGCTGCTGGACGCAACAGAGAGGATATGCAATTTG 717
Db 61 ACCAAAGTGGACAAGGTGTTGCTGCTGGACGCAACAGAGGATATGCAATTTG 120
Qy 718 GTTGTAGGCTTAATGACACATGAGAAATCTCTTGCTGTGTGACAGAAATGAGGCT 777

Db 121 GTTGTAGGCTTAATGACACATGAGAAATCTCTTGCTGTGTGACAGAAATGAGGCT 180
Qy 778 GAGATTTCTCTTCCACCTTGTATGCCATTCGCTGTGATGAGAAATGTTCTTTCAAT 837
Db 181 GAGATTTCTCTTCCACCTTGTATGCCATTCGCTGTGATGAGAAATGTTCTTTCAAT 240
Qy 838 AATGAAACCCCTCAGAAACATTTTGTACCAAGGCTGATTTGATTTGCTCCATCGCAGAAC 897
Db 241 AATGAAACCCCTCAGAAACATTTTGTACCAAGGCTGATTTGATTTGCTCCATCGCAGAAC 300
Qy 898 ACTTTGATTTGGTGGAGATGACTTCAAGAGTGTGACACCAAAATGAATCTGTGTGGTT 957
Db 301 ACTTTGATTTGGTGGAGATGACTTCAAGAGTGTGACACCAAAATGAATCTGTGTGGTT 360
Qy 958 GATTTCTTTGGGGGGCGTGTATCAAGGCAACATCATATAGTTAGTTAACAACATCGGGA 1017
Db 361 GATTTCTTTGGGGGGCGTGTATCAAGGCAACATCATATAGTTAGTTAACAACATCGGGA 420
Qy 1018 AACATGATGATGATGATATCTCTCGGCTCCAGAAACCTTCGCTCCAGAAATCTCCAG 1077
Db 421 AACATGATGATGATGATATCTCTCGGCTCCAGAAACCTTCGCTCCAGAAATCTCCAG 480
Qy 1078 AGCAACGTTGTTGACGATATGTTGTCACAGCAATGTCATCTGTATGAGCTGTGTAACAT 1137
Db 481 AGCAACGTTGTTGACGATATGTTGTCACAGCAATGTCATCTGTATGAGCTGTGTAACAT 540

RESULT 5
A1442485
LOCUS
DEFINITION
sa22f05.y1 Gm-cl004 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl004-538 5' similar to SW:INOI CITPA P42802
MYO-INSITOL-1-PHOSPHATE SYNTHASE ;, mRNA sequence.

ACCESSION A1442485
VERSION A1442485.1 GI:4296509
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE
AUTHORS Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Other_ESTs: AW348136 corresponding to Gm-r1021-68 (3')
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. Possible reversed clone: similarity on wrong strand This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Insert Length: 1557 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 390
POLYA=No.

FEATURES
source
1..561
/organism="Glycine max"
/mol_type="mRNA"

Db	4	TGAATCTCTGGGCTCCACAAACCTTCGCGCTCCAGGAAATCTCCAGGACAAGTTGTTG	63
QY	1091	ACGATATGCTCAACAGCAATGCCATCTCTATAGACCTGTGAAATCTCCGACCATGTGG	1150
Db	64	ACGATATGCTCAACAGCAATGCCATCTCTATAGACCTGTGAAATCTCCGACCATGTGG	123
QY	1151	TTGTATTAATATGCTGCCCTTAACGTAGGGGATACGAAGAGCCATGGAATGATACACTT	1210
Db	124	TTGTATTAATATGCTGCCCTTAACGTAGGGGATACGAAGAGCCATGGAATGATACACTT	183
QY	1211	CAGAGATATTCATGGGTGGGAAAGAACACATTTGTTTGGCAACAACATGTGAGATTCCC	1270
Db	184	CAGAGATATTCATGGGTGGGAAAGAACACATTTGTTTGGCAACAACATGTGAGATTCCC	243
QY	1271	TTTTAGCTGCTCCTATTATCTTGAATTGGACTTGGTCTTTCTTGTGAGCTGAGCACTAGATCC	1330
Db	244	TTTTAGCTGCTCCTATTATCTTGAATTGGACTTGGTCTTTCTTGTGAGCTGAGCACTAGATCC	303
QY	1331	AGTTTAAAGCTGAAAATGAGGGAATAATCCACTAATCCACCGAGTGGCTACCAATTCGA	1390
Db	304	AGTTTAAAGCTGAAAATGAGGGAATAATCCACTAATCCACCGAGTGGCTACCAATTCGA	363
QY	1391	GCTATCTGACCAAGGCTCTCTGGTCCACCGGGTACACAGTGGTGAATCATTTGTCAA	1450
Db	364	GCTATCTGACCAAGGCTCTCTGGTCCACCGGGTACACAGTGGTGAATCATTTGTCAA	423
QY	1451	AGCAGCGTGCATGCTGGAAAAATATATGAGGGCTTGTGTGGATTGGCCCCAGAGAATA	1510
Db	424	AGCAGCGTGCATGCTGGAAAAATATATGAGGGCTTGTGTGGATTGGCCCCAGAGAATA	483
QY	1511	ACATGATTTCTGAGTACAAAGTGA 1533	
Db	484	ACATGATTTCTGAGTACAAAGTGA 506	

```

source
1. .541
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1066-4644"
/tissue_type="Leaf and shoot tip, salt stressed, 2 week old seedling"
/lab_host="DH10B"
/clone_lib="Gm-c1066"
/note="Vector: pBluescript II SK+; Site.1: EcoRI, Site.2: XhoI; The cDNA library was constructed from mRNA isolated from unexpanded leaves and the shoot tips of 2 week old seedling from the cultivar Williams. The 2 week old seedlings were salt stressed in a solution of 500mM NaCl for 3 days prior to harvesting. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dt) sequence with a XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

```

RESULT 7	
EM085458	
LOCUS	
DEFINITION	BM085458 541 bp mRNA linear EST 23-JUL-2004 ga3j7a06.y1 Gm-cl066 glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl066-4644.5' similar to TR:Q9SSVA Q9SSV4
ACCESSION	M50-Q9SSITOL-1-PHOSPHATE SYNTHASE., mRNA sequence.
VERSION	BM085458
KEYWORDS	BM085458.1 GI:16996086
SOURCE	EST.
ORGANISM	Glycine max (soybean)
REFERENCE	Glycine max
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; glycine. 1 (bases 1 to 541) Shoenmaker,R., Kelm,P., Vocklin,L., Erpelting,J., Correll,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Boveri,Y., Pearson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schumk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R. Public Soybean EST Project Unpublished (1999) Contact: Shoemaker R/Public soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel.: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163, email: info@biogeneticservices.com)
JOURNAL	
COMMENT	
TITLE	
FEATURES	High quality sequence stop: 426. Location/Qualifiers

Query Match	32.4%	Score 496	DB 4	Length 541
Best Local Similarity	100.0%	Pred. No. 6.2e-273	Indels 0	Gaps 0
Matches 496	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1038	CTCGGCTCCACAAACCTTCGGCTCCAGGAATCTCCAGAGCAACGTTGTTCAGTAT	1097	
Db	1	CTGGGCTCCACAAACCTTCGGCTCCAGGAATCTCCAGAGCAACGTTGTTCAGTAT	60	
QY	1098	GGTCAACAGCANTGCCATCTCTATAGAGCTGGTGAACATCCGACCATGTTGTTAT	1157	
Db	61	GGTCAACAGCANTGCCATCTCTATAGAGCTGGTGAACATCCGACCATGTTGTTAT	120	
QY	1158	TAACTATGTCCTTACGTAGGGGGATACCAAGAGCGCATGGATGATGACATTCAGAGAT	1217	
Db	121	TAACTATGTCCTTACGTAGGGGGATACCAAGAGCGCATGGATGATGACATTCAGAGAT	180	
QY	1218	ATTGATGGGTGGAAGAACAACATTTGTTTGCACACACATGTGAGGATTCCTTTTACG	1277	
Db	181	ATTGATGGGTGGAAGAACAACATTTGTTTGCACACACATGTGAGGATTCCTTTTACG	240	
QY	1278	TGCTCTATTATCTTGGACTTTGGTCTTTCTTGCTGAGCTGAGCATGAAATCCAGTTAA	1337	
Db	241	TGCTCTATTATCTTGGACTTTGGTCTTTCTTGCTGAGCTGAGCATGAAATCCAGTTAA	300	
QY	1338	AGCTGAAATGAGGGGAAATTCACACTATTCACACCGAGTGTACCATTCACAGCTATCT	1397	
Db	301	AGCTGAAATGAGGGGAAATTCACACTATTCACACCGAGTGTACCATTCACAGCTATCT	360	
QY	1398	GACCAAGGCTCTCTGTGTTCCACCGGGTACACAGTGTGAATGATGTCAAAGCAGC	1457	
Db	361	GACCAAGGCTCTCTGTGTTCCACCGGGTACACAGTGTGAATGATGTCAAAGCAGC	420	
QY	1458	TGCATGCTGGAAAACATAATGAGGGCTTGTGTGATTTGGCCCCAGAAATAACATGAT	1517	
Db	421	TGCATGCTGGAAAACATAATGAGGGCTTGTGTGATTTGGCCCCAGAAATAACATGAT	480	
QY	1518	TCTGAGTACAAAGTGA 1533		
Db	481	TCTGAGTACAAAGTGA 496		

RESULT 8

AM348136/c

LOCUS

DEFINITION

782 bp

max cdna

Gm-r1021

Glycine

EST 04-OCT-2000

Gm-r1021-68 3', mRNA

sequence.

LOCUS	DEFINITION	LOCUS	DEFINITION
1175	TAGGGGATTAGCAAGAGAGCCATGGATAGTACATTCGAGAGATATTTCATGGGTGGAAAAGA	1234	
522	TAGGGGATTAGCAAGAGAGCCATGGATAGTACATTCGAGAGATATTTCATGGGTGGAAAAGA	463	
1235	ACAACCATTTGTTTTTGCAACAACAATGTGAGGATTCCTTTTAGCTGCTCTATTTATCTTTGG	1294	
462	ACAACCATTTGTTTTTGCAACAACAATGTGAGGATTCCTTTTAGCTGCTCTATTTATCTTTGG	403	
1295	ACCTTGATCTCTTTCTTGCTGAGCTGAGCCTTAGAATCCGATTTAAAGCTGAAAAATGAGGGA	1354	
402	ACTTGGTCTCTTTCTTGCTGAGCTGAGCCTTAGAATCCGATTTAAAGCTGAAAAATGAGGGA	343	
1355	AATTCCACTATTTCCACCCAGTGTGCTACATTTCTCAGCTATTTCTGACCAAGGCTCTCTG	1414	
342	AATTCCACTATTTCCACCCAGTGTGCTACATTTCTCAGCTATTTCTGACCAAGGCTCTCTG	283	
1415	TTCCACCGGGGTACACACAGTGTGATCATTTGTCAAAAGCAGCGTGCATATGCTGGAAAACA	1474	
282	TTCCACCGGGGTACACACAGTGTGATCATTTGTCAAAAGCAGCGTGCATATGCTGGAAAACA	223	
1475	TAATGAGGGCTTGTTGTTGATTTGGCCCCCAGAGATAATCATGATTTCTGAGTACCAAGTGA	1533	
222	TAATGAGGGCTTGTTGTTGATTTGGCCCCCAGAGATAATCATGATTTCTGAGTACCAAGTGA	164	
RESULT 9	BG239011		
LOCUS	BG239011	585 bp	mRNA
DEFINITION	BG239011	linear	EST 23-JUN-2004
	BB57809.y1 Gm-cl043 Glycine max cDNA clone GENOME SYSTEMS CLONE		
	ID: Gm-cl043-3258 5' similar to TR:Q9SSV4 Q9SSV4		
	MTO-INOSITOL-1-PHOSPHATE SYNTHASE. ;, mRNA sequence.		

REFERENCE
1 (bases 1 to 585)
Shoemaker, R., Keim, P., Vodkin, L., Expediting, J., Corvelli, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucab, T., Martin, J.,
Beck, C., White, T., Underwood, K., Septece, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N.,
Schurr, R., Rittner, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
TITLE
JOURNAL
COMMENT
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel.: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)

```

FEATURES
    source
        Location/Qualifiers
            1..585
                /organism="Glycine max"
                /mol_type="mRNA"
                /cultivar="Williams"
                /db_xref="taxon:3847"
                /clone="GENOME SYSTEMS CLONE ID: Gm-c1043-3258"
                /tissue_type="Hypocotyl and Plumule, germinating seeds"
                /lab_host="DH10B"
                /clone_1fb="Gm-c1043"
                /note="Vector: pRT3pac (pharmacia); Site_1: EcoRI;

```

Site 2: NtII. This cDNA library was constructed from mRNA isolated from hypocotyl and plumule tissues of seeds germinated for three days of the cultivar Williams. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a NotI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and NotI. The cDNA fragments were directionally cloned into the EcoRI-NotI restriction site of the pT73-Pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."

ORIGIN

Query Match	30.4%;	Score 466;	DB 4;	Length 585;
Best Local Similarity	100.0%;	Pred. No. 1.2e-255;		
Matches 466;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	893	GGAAACCTTTGATTGGTGGAGATAGATCTCAAGAGTGGTCAACCAAAATGAAATCTGTGT	952
Db	14	GGAAACCTTTGATTGGTGGAGATAGATCTCAAGATGGTCAACCAAAATGAAATCTGTGT	73
QY	953	TGTTGATTTTCTTGTGGGGCTGGTATCAAGCCAATCTATAGTTAGTTACAACATC	1012
Db	74	TGTTGATTTTCTTGTGGGGCTGGTATCAAGCCAATCTATAGTTAGTTACAACATC	133
QY	1013	TGGGAAACAATGATGATGATGAAATCTCTCGGTCCACAAACCTTCCGCTCCAGGAATCT	1072
Db	134	TGGGAAACAATGATGATGATGAAATCTCTCGGTCCACAAACCTTCCGCTCCAGGAATCT	193
QY	1073	CCAAGACCAACGTGTGACGATATGTGCACACAGCAATGCCATCCTATAGGCTGGTG	1132
Db	194	CCAAGACCAACGTGTGACGATATGTGCACACAGCAATGCCATCCTATAGGCTGGTG	253
QY	1133	AACATCCCGACCATGTGTTGTTATTAAGTATGTGCTTACGTAGGGGATAGCAAGAG	1192
Db	254	AACATCCCGACCATGTGTTGTTATTAAGTATGTGCTTACGTAGGGGATAGCAAGAG	313
QY	1193	CCATGATGATGATACCTCAGAGATATTCATGGGTGGAAGAAACAACCATTTGTTTGACA	1252
Db	314	CCATGATGATGATACCTCAGAGATATTCATGGGTGGAAGAAACAACCATTTGTTTGACA	373
QY	1253	ACACATGTGAGGATTCCTTTTAACTGCTCCTATATATCTTGGACTTGTGCTTCTGTG	1312
Db	374	ACACATGTGAGGATTCCTTTTAACTGCTCCTATATATCTTGGACTTGTGCTTCTGTG	433
QY	1313	AGCTGAGACCTAGAAATCCAGTTTAAAGTGAAGAAATGAGGAAAAATT	1358
Db	434	AGCTGAGACCTAGAAATCCAGTTTAAAGTGAAGAAATGAGGAAAAATT	479

ORIGIN

Query Match	30.1%	Score 462;	DB 2;	Length 618;
Best Local Similarity	-99.8%	Pred. No. 2.3e-253;		
Matches 512; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	261	CATTTATGGGCTACAAAGAGAAABAATTCACAAGGCCAATACCTTGGCTCCCGACCCA	320
Db	45	CATTTATGGGCTACAAAGAGAAABAATTCACAAGGCCAATACCTTGGCTCCCGACCCA	104
Qy	321	AGGCTCAGCATTTCCGATGGATTCCTTCGAGGAGAGAAATCTATGGCCCATTCAGAG	380
Db	105	AGGCTCAGCATTTCCGATGGATTCCTTCGAGGAGAGAAATCTATGGCCCATTCAGAG	160
Qy	381	TCGCTTCCAAATGGTTAATCTCGACGACATGTGTTTGGGGATATGGAATACAGCAACAT	440
Db	165	TCGCTTCCAAATGGTTAATCTCGACGACATGTGTTTGGGGATATGGAATACAGCAACAT	224
Qy	441	GAACCTGGCTGATGCCATGGCCAGGGCAAAAGTGTTCATCGATTTGCAGAAAGCAGTT	500
Db	225	GAACCTGGCTGATGCCATGGCCAGGGCAAAAGTGTTCATCGATTTGCAGAAAGCAGTT	284
Qy	501	GAGGCGCTTACATGGAATCCATGGTTCACATCCCGGAATCTACGACCCGGATTTCAATTGC	560
Db	285	GAGGCGCTTACATGGAATCCATGGTTCACATCCCGGAATCTACGACCCGGATTTCAATTGC	344
Qy	561	TGCGAACCAAGAGAGAGCGTGCCACAACAGTGTATTAGGGCCAAAGCAAGCAAGCAATTCA	620
Db	345	TGCGAACCAAGAGAGAGCGTGCCACAACAGTGTATTAGGGCCAAAGCAAGCAAGCAATTCA	404
Qy	621	GCAATATCATAAAGACATCAAGGCGTTTAAAGAAAGCCCAAAATGGACAAAGTGGTGT	680

Db	405	GCAATCATCAAAGCATCAAGCCCTTTAAAGAACCCAAAGGTGACAAGGTTGTTGT	464
Qy	681	CCTGNGCTGSCCAACAGAGAGGTATAGCAATTGTGTGAGGCTTAATGACACAT	740
Db	465	CTGTGACCTGCCAAACAGAGAGGTATAGCAATTGTGTGAGGCTTAATGACACAT	524
Qy	741	GGAGAAATCTCTTGGCTGCTGTGGACACAGAAATGA	773
Db	525	GGAGAAATCTCTTGGCTGCTGTGGACACAGAAATGA	557

RESULT 11	LOCUS	DEFINITION	LOCUS	DEFINITION	LOCUS	DEFINITION
AM203517	AM203517	449 bp	mRNA	linear	EST 16-JUL-2004	
g53b05.y1	Gm-c1028	Glycine max	cDNA	clone	GENOME SYSTEMS	CLONE ID: ID
Gm-c1028-1930	5'	similar to	SM:IN01	SP10	P42803	
MYO-INOSITOL-1-PHOSPHATE SYNTHASE	;	mRNA	sequence.			

ACCESSION	AM20351/	GI:502144
VERSION	AM203517.1	
KEYWORDS	EST.	
SOURCE	Glycine max (soybean)	
ORGANISM	Glycine max	

REFERENCE
AUTHORS

1 (bases 1 to 449)
Shoemaker, R., Kelm, P., Vodka, L., Eipel, J., Coryell, V.,
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

TITLE Public Soybean EST Project
JOURNAL R/Public Soybean EST Project
COMMENT Unpublished (1999)
Contact: Shoemaker

Public Soybean EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Insert length: 1185 Std Error: 0.00
High quality sequence stop: 416.

```

Location/Qualifiers
1. .449
/origanism="Glycine max"
/mol_type="mRNA"
/cultivar="Supernod"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1028-1930"
/tissue_type="roots of 'Supernod' plants"
/lab_host="DH10B"
/clone_id="Gm-c1028"
/notes="Vector: pBluescript II XR; Site 1: EcoRI; Site_2:
XhoI; The mRNA was isolated from roots of Glycine max
'Supernod' plants generously donated by Dr. Gary Stracey.
The seedlings were inoculated with Bradyrhizobium
japonicus, strain USDA110 prior to harvest. Stracey's
cDNA synthesis kit (catalog number 200401) was used to
synthesize the cDNA. First-strand synthesis was performed
with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stracey's
first-strand synthesis primer was used. An 'anchor'
nucleotide (V=A,C, or G) was added to the 3' end of the
primer (GAGACGAGACGAGACGAGACGAGTCGAG(T)18V) to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in with
Cloned Pfu DNA polymerase, ligated to EcoRI adapters and

```

```

ORIGIN
Query Match      29.0%; Score 444; DB 2; Length 449;
Best Local Similarity 100.0%; Pred. No. 5.1e-243;
Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500BP cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript II XR predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Corryell."

QY	634	GACATCAAGCGTTTAAGGAAGCACCAAAAGTGACAAGGTGGTTCGTGTGGCTGCG	693
Db	6	GACATCAAGCGTTTAAGGAAGCACCAAAAGTGACAAGGTGGTTCGTGTGGCTGCG	65
QY	694	AACACAGAGAGGTATAGCAATTGGTGTGAGGCTTATATGACACATGAGAAATCTCTTG	753
Db	66	AACACAGAGAGGTATAGCAAAATGGTGTGAGGCTTATATGACACATGAGAAATCTCTTG	125
QY	754	GCTGCTGTGACAGAAATGAGGCGTGAGATTCTCTCCACCTTGATAGCATTTGCCGTG	813
Db	126	GCTGCTGTGACAGAAATGAGGCGTGAGATTCTCTCCACCTTGATAGCATTTGCCGTG	185
QY	814	GTCATGAGAAATGTTCCCTTCATTATATGAAGCCCTCAGAACATTTTGTATCAAGGCGTG	873
Db	186	GTCATGAGAAATGTTCCCTTCATTATATGAAGCCCTCAGAACATTTTGTATCAAGGCGTG	245
QY	874	ATTGATTTTGGCCATCGGAGAAACATTTGATTTGGTGTGAATGATCTTCAGAGTGTCTAG	933
Db	246	ATTGATTTTGGCCATCGGAGAAACATTTGATTTGGTGTGAATGATCTTCAGAGTGTCTAG	305
QY	934	ACCAAAATGAAATCTGTGTGGTGTGATTTTCTTGTGGGGCGCTGTATCAAGCCAACTCT	993
Db	306	ACCAAAATGAAATCTGTGTGGTGTGATTTTCTTGTGGGGCGCTGTATCAAGCCAACTCT	365
QY	994	ATAGTTAGTTACAAACCATCTGGGAAACAAATGATGTATGAAATCTCTGGCTCCAAACC	1055
Db	366	ATAGTTAGTTACAAACCATCTGGGAAACAAATGATGTATGAAATCTCTGGCTCCAAACC	425
QY	1054	TTCCGGCTCCAAAGAAATCTCCAG 1077	
Db	426	TTCCGGCTCCAAAGAAATCTCCAG 449	

RESULT 12	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
CA938527		632 bp mRNA linear EST 01-JUL-2004	CA938527						
		seq34a11.y1 Gm-cl023 glycine max cDNA clone SOYBEAN							
		Gm-cl023-6165.5, similar to TR:09SSV4 09SSV4							
		MYO-INOSITOL-1-PHOSPHATE SYNTHASE. ;, mRNA sequence.							
			CA938527						
			CA938527.1	GI:27427007					
		EST.							
		glycine max (soybean)							
		glycine max							
		Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;							
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;							
		rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;							
		Glycine							
		1 (bases 1 to 632)							
		Shoenmaker,R., Keim,P., Vodkin,L., Eppelding,J., Coryell,V.,							
		Khanma,A., Bolla,B., Marra,M., Hiller,L., Kucaba,T., Martin,J.,							
		Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,							
		Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,							

Schurk, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCam, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 447.

FEATURES

Location/Qualifiers

1..632
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="1157"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl023-6165"
/tissue_type="seed coats of greenhouse grown plants"
/lab_host="DH10B"
/clone_lib="Gm-cl023"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from seed
coats (100-200mgs) of greenhouse grown plants. The
library was prepared using the Life Technologies
pSuperScript cDNA library construction kit. Complementary
DNA was synthesized from mRNA using a poly (dT) sequence
with a Not I restriction site. Sal I linkers adapters
were ligated to the blunt-ended cDNA fragments followed by
Not I digestion. The cDNA fragments were directionally
cloned into the Not I-Sal I restriction site of the
pSPORT1 vector. The ligated cDNA fragments were
transformed into E.coli Electromax DH10B host cells. This
library was constructed by Dr. Lila Vodkin and Dr. Anu
Khanna."

ORIGIN

Query Match 27.3%; Score 419; DB 6; Length 632;
Best Local Similarity 99.8%; Pred. No. 1.3e-228;
Matches 469; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
1 ATGTTTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
111 ATGTTTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 170
61 CAGTCGGTGTACAGCAACCAACCAACCACTGTTCAAGAGACAGAAAGGCACTGAT 120
171 CAGTCGGTGTACAGCAACCAACCAACCACTGTTCAAGAGACAGAAAGGCACTGAT 230
121 CAGTCGGTGTACAGCAACCAACCAACCACTGTTCAAGAGACAGAAAGGCACTGAT 180
231 CAGTCGGTGTACAGCAACCAACCAACCACTGTTCAAGAGACAGAAAGGCACTGAT 290
181 AAATTTGGGGGTGATGTTGTGGGTGGGGTGAACAACAGGCTTACCGTGGT 240
291 AAATTTGGGGGTGATGTTGTGGGTGGGGTGAACAACAGGCTTACCGTGGT 350
241 GTTATTTGCTAACAGAGAGACATTTTCATGGGCTACAAAGAGACAGATTTCAACAGCAAT 300
351 GTTATTTGCTAACAGAGAGACATTTTCATGGGCTACAAAGAGACAGATTTCAACAGCAAT 410
301 TACTTTGGCTCCCTCAACCCAGCCTCAGCTATTTCGAGTTGAGTCTTCCAGGAGAGGAA 360
411 TACTTTGGCTCCCTCAACCCAGCCTCAGCTATTTCGAGTTGAGTCTTCCAGGAGAGGAA 470
361 ATCTATGCCCATTTCAAGAGTCTGCTTCAATGGTTAATCTCGAGACATTTGTGTTGGG 420
```

Db 471 ATCTATGCCCATTTCAAGAGTCTGCTTCAATGGTTAATCTCGAGACATTTGTGTTGGG 530
Qy 421 GGATGGATATACGACATGACATGACCTGCTGATGCGATGCGAGGCAAA 470
Db 531 GGATGGATATACGACATGACATGACCTGCTGATGCGATGCGAGGCAAA 580

RESUT 13
Bg652636
LOCUS
DEFINITION
Bg652636 516 bp mRNA linear EST 22-JUL-2004
ID: Gm-cl051-5668 5' similar to TR:Q9SSV4 Q9SSV4
MYO-INOSITOL-1-PHOSPHATE SYNTHASE.; mRNA sequence.

ACCESSION

VERSION Bg652636.1 GI:13790045

KEYWORDS

SOURCE

ORGANISM Glycine max (soybean)

Glycine max (soybean)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

1 (bases 1 to 516)
Shoemaker, R., Kaim, P., Vodkin, L., Expelding, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
McCam, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCam, R., Waterston, R. and Wilson, R.

REFERENCE

AUTHORS

Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

JOURNAL

COMMENT

When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
High quality sequence stop: 383.

FEATURES

source

1..516
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Corolla"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-5668"
/tissue_type="floral meristematic mRNA"
/lab_host="DH10B"
/clone_lib="Gm-cl051"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from floral
meristematic mRNA provided by Dr. Halina Knop of Clemson
University. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."

ORIGIN

Query Match 26.9%; Score 413; DB 4; Length 516;
Best Local Similarity 99.6%; Pred. No. 3.6e-225;
Matches 513; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
812 GTGTGATGGAATATGTTCTTTCATTTATGGAAGGCTTCAACACTTTTGTACAGGCG 871
```

Db 2 GTGTGATGAAAAATGTTCTTTCATTATGAAAGCCCTCAGAACCTTTTGTACAGAGGC 61
 QY 872 TGAATTATCTTGGCATTCGGAGAGACACTTGTGATGAGATGACTTCAAGATGCTC 931
 Db 62 TGAATTATCTTGGCATTCGGAGAGACACTTGTGATGAGATGACTTCAAGATGCTC 121
 QY 932 AGACCAAAATGAAATCTGTGTGATTTCTGTGTGGGGGCTGGATCAAGCCAAACAT 991
 Db 122 AGACCAAAATGAAATCTGTGTGATTTCTGTGTGGGGGCTGGATCAAGCCAAACAT 181
 QY 992 CTATAGTTAGTTACCAACATCTGGGAAACATGATGATGATGATCTCTCGCTCCACAAA 1051
 Db 182 CTATAGTTAGTTACCAACATCTGGGAAACATGATGATGATGATCTCTCGCTCCACAAA 241
 QY 1052 CCTTCCGCTCCAGAGAAATCTCCAGAGCAACTTGTGACGATGATGATGATGATG 1111
 Db 242 CCTTCCGCTCCAGAGAAATCTCCAGAGCAACTTGTGACGATGATGATGATGATG 301
 QY 1112 CCATCTCTATAGAGCTGGGTGAACATCCGACCATGTGTGATTTAGTATGATGCTT 1171
 Db 302 CCATCTCTATAGAGCTGGGTGAACATCCGACCATGTGTGATTTAGTATGATGCTT 361
 QY 1172 AGCTAGGGGATGACAGAGAGAGCCATGATGATGATGATGATGATGATGATGATG 1231
 Db 362 AGCTAGGGGATGACAGAGAGAGCCATGATGATGATGATGATGATGATGATGATG 421
 QY 1232 AGAACCACTTGTGTGACACACATGTGAGAGATTCCTTTAGTGTCTCTATTATCT 1291
 Db 422 AGAACCACTTGTGTGACACACATGTGAGAGATTCCTTTAGTGTCTCTATTATCT 481
 QY 1292 TGGACTTGGTCTCTTGTGCTGAGCTGAGCACTGA 1326
 Db 482 TGGACTTGGTCTCTTGTGCTGAGCTGAGCACTGA 516

RESULT 14
 BE660322 505 bp mRNA linear EST 06-SEP-2000
 LOCUS 1274 GmaxSC Glycine max cDNA, mRNA sequence.
 DEFINITION BE660322
 VERSION BE660322.1 GI:9986214
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 505)
 HARRIS, N., CHAPMAN, B. P. and GILZEN, M.
 Gene expression in developing soybean seed coats
 Unpublished (2000)
 CONTACT: Gilzen M
 Agriculture and Agri-Food Canada
 1391 Sandford Street, London, Ontario, Canada N5V 4T3
 Tel: 519 457 1470
 Fax: 519 457 3997
 Email: gilzenm@agr.gc.ca.
 Location/Qualifiers
 1..505
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Harosoy 63"
 /db_xref="taxon:3847"
 /tissue_type="Seed coats"
 /lab_host="E. coli strain XL0LR"
 /note="Vector: pBK-CMV; Site 1: EcoRI, Site 2: XhoI. This
 cDNA library was constructed from polyA+ enriched mRNA
 from green seed coats in mid to late developmental
 stage, average fresh weight 250 mg per seed. Traces of
 pod and embryo tissue also present. Complementary DNA was
 synthesized from mRNA using an XhoI-poly(dT)

linker-primer. EcoRI adapters were ligated to the
 blunt-ended cDNA fragments and the products were digested
 with XhoI for directional cloning into lambda ZAP Express
 vector. This lambda library was amplified once using E.
 coli host strain XL1 Blue MRF+. Inserts were then
 subcloned by mass excision using ExbaSst helper phage for
 conversion into phagemid vector pBK-CMV in E. coli host
 strain XL0LR."

ORIGIN

Query Match 24.6%; Score 377; DB 2; Length 505;
 Best local Similarity 99.8%; Pred. No. 1,9e-204;
 Matches 427; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTCAATGAGAAATTTTAAGGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
 Db 77 ATGTTCAATGAGAAATTTTAAGGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 136
 QY 61 CAGTCGGGTACAACTAGCAAAACCAACCGAATCTGTTCAAGAGAACAGAAATGGCACTTAT 120
 Db 137 CAGTCGGGTACAACTAGCAAAACCAACCGAATCTGTTCAAGAGAACAGAAATGGCACTTAT 196
 QY 121 CAGTGATTTGTCAAAACCAATCCGTCACATCAATTTAAACCAACCAACCAATGTTCCA 180
 Db 197 CAGTGATTTGTCAAAACCAATCCGTCACATCAATTTAAACCAACCAACCAATGTTCCA 256
 QY 181 AAATTGGGGGTGATGCTGTGGGTGGGGTGAACCAACGCTCTTACCTCAACCGGTGCT 240
 Db 257 AAATTGGGGGTGATGCTGTGGGTGGGGTGAACCAACGCTCTTACCTCAACCGGTGCT 316
 QY 241 GTTATTGCTCAAGAGAGAGACATTTTCATGAGGCTTACAAAGAGAAAGATTAACAAGCAAT 300
 Db 317 GTTATTGCTCAAGAGAGAGACATTTTCATGAGGCTTACAAAGAGAAAGATTAACAAGCAAT 376
 QY 301 TACTTGGCTCCCTCAACCAACCTCAGCTATTCAGATGATGATGATGATGATGATGATG 360
 Db 377 TACTTGGCTCCCTCAACCAACCTCAGCTATTCAGATGATGATGATGATGATGATGATG 436
 QY 361 ATCTATGCCCATTCAGAGAGTCTGCTTCCATGTTAATCTGACGACCAATGTGTTGGG 420
 Db 437 ATCTATGCCCATTCAGAGAGTCTGCTTCCATGTTAATCTGACGACCAATGTGTTGGG 496
 QY 421 GGATGGGA 428
 Db 497 GGATGGGA 504

RESULT 15
 BUS49101/c 588 bp mRNA linear EST 16-SEP-2002
 LOCUS GM880018B10C10 Gm-r1088 Glycine max cDNA clone Gm-r1088-6667 3',
 DEFINITION BUS49101
 mRNA sequence.
 ACCESSION BUS49101
 VERSION BUS49101.1 GI:22931962
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

1 (bases 1 to 588)
 VODKIN, L., SHOEMAKER, R., KEIM, P., RETZEL, E., KHANNA, A., SHEALY, R.,
 CLOUGH, S., THIBAUD-MISEN, F., CORYELL, V., ERPELDING, J., RAPH, C.,
 SHOOP, E., STROMVIK, M., SCHWEITZER, P., GONG, G. and LIU, L.
 A Functional Genomics Program for Soybean (NSF 9872565) (2002)
 Unpublished (2002)
 Other ESTs: BF324641 corresponding to Gm-c1066-1370 (5')
 Contact: Vodkin, L.O., PI, A Functional Genomics Program for
 Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional
 Genomics
 University of Illinois

Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l-vodkin@uiuc.edu
 Insert Length: 588 Std Error: 0.00
 Plate: GM880018B10 row: C column: 10
 Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'
 High quality sequence stop: 588.
 Location/Qualifiers
 1. 588
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="caxon:3847"
 /clone_lib="Gm-r1088-6667"
 /clone_lib="Gm-r1088"
 /note="The library Gm-r1088 is a sequence-driven, reracked set of 9,216 low redundancy clones selected from cDNA libraries from various tissues and stages of development of soybean that consists of 2,706 cDNAs from germinating cotyledons (source library Gm-c1027); 1,355 cDNAs from immature seed coats (libraries Gm-c1019 and Gm-c1023); 917 cDNAs from tissue culture derived somatic embryos (source libraries Gm-c1036 and Gm-c1075); 3,275 cDNAs from germinating seedlings, shoot tips, or leaves exposed to various stresses (source libraries Gm-c1064, Gm-c1065, Gm-c1066, and Gm-c1067; and Gm-c1068); and 963 cDNAs from young leaves exposed to bacterial and fungal pathogens (source libraries Gm-c1072, Gm-c1073, and Gm-c1074). The 5' ESTs of the source clones from the different libraries was used to select singletons, or a representative of each contig, which were reracked to form library Gm-r1088 and the cDNA clones of the reracked Gm-r1088 library were then sequenced at the 3' end. The unigene selection and 3' sequencing was funded by NSF Plant Genome project #9872565 (http://soybean.genomics.croptsci.uiuc.edu/) as part of creation of a low redundancy soybean cDNA set. The source cDNA libraries were constructed by the laboratories of Lila Vodkin, University of Illinois, Randy Shoemaker, Iowa State University, and Paul Keim, Northern Arizona University as part of the Public EST project. http://129.186.26.94/soybeanest.html. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota, http://web.ahc.umn.edu/biodata/nsfsoy/. Reracking and 3' sequencing were conducted by services of the University of Illinois Keck Center for Comparative and Functional Genomics http://www.biotech.uiuc.edu/keck.htm. Note: The corresponding 5' EST from each clone in the Gm-r1088 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Genome Systems clone ID of the original cDNA library that is also listed under 'OTHER EST'."

ORIGIN

Query Match 24.4%; Score 374; DB 5; Length 588;
 Best Local Similarity 100.0%; Pred. No. 1e-202;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1160 AGTATGTCCTTACGTAGGAGATAGCAAGAGCCATGATGATACCTTACAGATAT 1219
 DB 501 AGTATGTCCTTACGTAGGAGATAGCAAGAGCCATGATGATACCTTACAGATAT 442
 QY 1220 TCATGGGTGAAAGAACACCATTTGTTTGCACAACACATGTGAGATTCCCTTTAGCTG 1279
 DB 441 TCATGGGTGAAAGAACACCATTTGTTTGCACAACACATGTGAGATTCCCTTTAGCTG 382
 QY 1280 CTCCTATTATCTTGGACTGGTCCCTCTGAGCTGAGCACTAGATCCAGTTAAAG 1339
 DB 381 CTCCTATTATCTTGGACTGGTCCCTCTGAGCTGAGCACTAGATCCAGTTAAAG 322
 QY 1340 CTGAAATGAGGAAATTCACCTCACCAGTTCCTACATTCACGTAATCTGA 1399

DB 321 CTGAAATGAGGAAATTCACCTCACCAGTTCCTACATTCACGTAATCTGA 262
 QY 1400 CCAAGGCTCCTCTGTTTCCACCGGGTACACAGTGTGATGATTTGCAAGACGCTG 1459
 DB 261 CCAAGGCTCCTCTGTTTCCACCGGGTACACAGTGTGATGATTTGCAAGACGCTG 202
 QY 1460 CAATGCTGAAACATTAATGAGGCTTGTGTTGATTTGGCCCAAGAAATACATGATTC 1519
 DB 201 CAATGCTGAAACATTAATGAGGCTTGTGTTGATTTGGCCCAAGAAATACATGATTC 142
 QY 1520 TCGAGTACAGTGA 1533
 DB 141 TCGAGTACAGTGA 128

Search completed: June 8, 2005, 14:33:48
 Job time : 525.88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 06:50:16 ; Search time 260.698 Seconds
(without alignments)
9621.891 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533
Sequence: 1 atgctcgcgcagatcttaaa.....tgattctgcgacgaagtgca 1533

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
6: /cgn2_6/prodata/1/ina/5B_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	23	1.5	1931 3	US-09-118-442-10 Sequence 10, Appl
2	23	1.5	1931 3	US-09-677-064-10 Sequence 10, Appl
3	23	1.5	3546 3	US-09-118-442-14 Sequence 14, Appl
4	23	1.5	3546 3	US-09-118-442-15 Sequence 15, Appl
5	23	1.5	3546 3	US-09-677-064-14 Sequence 14, Appl
6	23	1.5	3546 3	US-09-677-064-15 Sequence 15, Appl
7	20	1.3	27223 4	US-09-949-016-13036 Sequence 13036, A
8	19	1.2	1258 1	US-08-539-798-1 Sequence 1, Appl
9	19	1.2	1258 1	US-08-329-560-1 Sequence 1, Appl
10	19	1.2	3153 3	US-09-175-928-9 Sequence 9, Appl
11	18	1.2	256 4	US-09-313-294A-2055 Sequence 2055, Ap
12	18	1.2	271 4	US-09-513-999C-9342 Sequence 9342, Ap
13	18	1.2	472 3	US-09-641-638-549 Sequence 549, App
14	18	1.2	472 3	US-09-641-638-550 Sequence 550, App
15	18	1.2	472 3	US-10-170-097-549 Sequence 549, App
16	18	1.2	472 3	US-10-170-097-550 Sequence 550, App
17	18	1.2	601 4	US-09-949-016-41180 Sequence 41180, A
18	18	1.2	601 4	US-09-949-016-41180 Sequence 41180, A
19	18	1.2	601 4	US-09-949-016-112271 Sequence 112271, Sequence 141642,
20	18	1.2	1001 3	US-09-641-638-278 Sequence 278, App
21	18	1.2	1001 3	US-10-170-097-278 Sequence 278, App
22	18	1.2	1398 4	US-09-328-352-429 Sequence 429, App
23	18	1.2	2477 4	US-09-907-794A-169 Sequence 169, App
24	18	1.2	2477 4	US-09-905-125A-169 Sequence 169, App
25	18	1.2	2477 4	US-09-902-775A-169 Sequence 169, App
26	18	1.2	2477 4	US-09-906-700-169 Sequence 169, App
27	18	1.2	2477 4	US-09-903-603A-169 Sequence 169, App

ALIGNMENTS

C 28	18	1.2	2477 4	US-09-904-920A-169	Sequence 169, App
C 29	18	1.2	2477 4	US-09-909-064-169	Sequence 169, App
C 30	18	1.2	2477 4	US-09-905-381A-169	Sequence 169, App
C 31	18	1.2	2477 4	US-09-906-618-169	Sequence 169, App
C 32	18	1.2	4989 4	US-09-693-011-12	Sequence 12, Appl
C 33	18	1.2	5083 4	US-09-693-011-11	Sequence 11, Appl
C 34	18	1.2	6314 4	US-09-693-011-10	Sequence 10, Appl
C 35	18	1.2	6408 4	US-09-693-011-9	Sequence 9, Appl
C 36	18	1.2	9060 3	US-08-378-313-20	Sequence 20, Appl
C 37	18	1.2	16573 4	US-09-949-016-11764	Sequence 11764, A
C 38	18	1.2	17370 4	US-09-949-016-17331	Sequence 17331, A
C 39	18	1.2	55264 4	US-09-949-016-15014	Sequence 15014, A
C 40	18	1.2	56147 4	US-09-949-016-16352	Sequence 16352, A
C 41	18	1.2	87617 4	US-09-949-016-16551	Sequence 16551, A
C 42	18	1.2	92227 4	US-09-949-016-11929	Sequence 11929, A
C 43	18	1.2	92232 4	US-09-949-016-15421	Sequence 15421, A
C 44	18	1.2	94755 4	US-09-949-016-11839	Sequence 11839, A
C 45	18	1.2	107140 4	US-09-949-016-14834	Sequence 14834, A

RESULT 1
US-09-118-442-10
Sequence 10, Application US/09118442B
Patent No. 6197561
GENERAL INFORMATION:
APPLICANT: Martino-Catt, Susan J.
APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.
APPLICANT: Wang, Xun
APPLICANT: Bowen, Benjamin A.
TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
FILE REFERENCE: 0706
CURRENT APPLICATION NUMBER: US/09/118,442B
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: 60/055,446
EARLIER FILING DATE: 1997-08-11
EARLIER APPLICATION NUMBER: 60/055,526
EARLIER FILING DATE: 1997-08-08
EARLIER APPLICATION NUMBER: 60/053,944
EARLIER FILING DATE: 1997-07-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 1931
TYPE: DNA
ORGANISM: Zea mays
US-09-118-442-10
Query Match 1.5%; Score 23; DB 3; Length 1931;
Best local Similarity 100.0%; Pred. No. 0.22;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1105 AGCAATGCCATCTCTATGAGCC 1127
Db 1203 AGCAATGCCATCTCTATGAGCC 1225
RESULT 2
US-09-677-064-10
Sequence 10, Application US/09677064
Patent No. 6291224
GENERAL INFORMATION:
APPLICANT: Martino-Catt, Susan J.
APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.
TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
FILE REFERENCE: 0706D
CURRENT APPLICATION NUMBER: US/09/677,064

```
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-10
```

```
Query Match 1.5%; Score 23; DB 3; Length 1931;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1105 AGCAATGCCATCCTCTATGAGCC 1127
DB 1203 AGCAATGCCATCCTCTATGAGCC 1225
```

```
RESULT 3
US-09-118-442-14
```

```
; Sequence 14, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phyate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-14
```

```
Query Match 1.5%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1105 AGCAATGCCATCCTCTATGAGCC 1127
DB 2842 AGCAATGCCATCCTCTATGAGCC 2864
```

```
RESULT 4
US-09-118-442-15
```

```
; Sequence 15, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
```

```
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phyate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-15
```

```
Query Match 1.5%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1105 AGCAATGCCATCCTCTATGAGCC 1127
DB 2842 AGCAATGCCATCCTCTATGAGCC 2864
```

```
RESULT 5
US-09-677-064-14
```

```
; Sequence 14, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phyate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-14
```

```
Query Match 1.5%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1105 AGCAATGCCATCCTCTATGAGCC 1127
DB 2842 AGCAATGCCATCCTCTATGAGCC 2864
```

```
RESULT 6
US-09-677-064-15
```

```
; Sequence 15, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
```

```

; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
; US-09-677-064-15

Query Match          1.5%; Score 23; DB 3; Length 3546;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1105 AGCAATGCCATCTCTATAGACC 1127
Db      2842 AGCAATGCCATCTCTATAGACC 2864
|||||
|||||

RESULT 7
US-09-949-016-13036
; Sequence 13036, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13036
; LENGTH: 27223
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-13036

Query Match          1.3%; Score 20; DB 4; Length 27223;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1476 AATGAGGCTGTGTGAT 1495
Db      3336 AATGAGGCTGTGTGAT 3355
|||||
|||||

RESULT 8
US-08-539-798-1
; Sequence 1, Application US/08539798
; Patent No. 5614400
; GENERAL INFORMATION:
; APPLICANT: CAHOON, Edgar B.
; APPLICANT: OHROGEE, John B.
; APPLICANT: OHROGEE, John B.
```

```

; TITLE OF INVENTION: Methods and Compositions Relating to
; TITLE OF INVENTION: Plant 6-Delta Palmitoyl-Acyl Carrier Protein Desaturase
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pioneer Hi-Bred International, Inc.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: US
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,798
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,560
; FILING DATE: 26-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Yates, Michael B.
; REGISTRATION NUMBER: 36,063
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1258 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-539-798-1

Query Match          1.2%; Score 19; DB 1; Length 1258;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      327 AGCTATTCGAGTTGGATCC 345
Db      819 AGCTATTCGAGTTGGATCC 837
|||||
|||||

RESULT 9
US-08-329-560-1
; Sequence 1, Application US/08329560
; Patent No. 5654402
; GENERAL INFORMATION:
; APPLICANT: CAHOON, Edgar B.
; APPLICANT: OHROGEE, John B.
; TITLE OF INVENTION: Methods and Compositions Relating to
; TITLE OF INVENTION: Plant 6-Delta Palmitoyl-Acyl Carrier Protein Desaturase
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pioneer Hi-Bred International, Inc.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: US
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/329,560
; FILING DATE: 26-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

NAME: Yates, Michael E.
REGISTRATION NUMBER: 36, 063
REFERENCE/DOCKET NUMBER: 0284US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1258 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-329-560-1

Query Match 1.2%; Score 19; DB 1; Length 1258;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 AGCTATTCGAGTTGGATCC 345
|||
Db 819 AGCTATTCGAGTTGGATCC 837

RESULT 10
US-09-175-928-9/c
Sequence 9, Application US/09175928A
Patent No. 6312921
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John W.
APPLICANT: Lavallee, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Mi, Sha
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: 6006B.AJ172A
CURRENT APPLICATION NUMBER: US/09/175,928A
CURRENT FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 62
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 3153
TYPE: DNA
ORGANISM: Homo sapiens
US-09-175-928-9

Query Match 1.2%; Score 19; DB 3; Length 3153;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1100 TCACAGCAATGCCATCCT 1118
|||
Db 178 TCACAGCAATGCCATCCT 160

RESULT 11
US-09-313-294A-2055
Sequence 2055, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Lalgudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 2055

LENGTH: 256
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6476212 700551970H1
US-09-313-294A-2055

Query Match 1.2%; Score 18; DB 4; Length 256;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 CAACATGAACCTGGCTGA 452
|||
Db 99 CAACATGAACCTGGCTGA 116

RESULT 12
US-09-513-999C-9342
Sequence 9342, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Mline Edwards, J.B.
APPLICANT: Duclet, A.Y.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
FILE REFERENCE: 59.US2.REG
CURRENT APPLICATION NUMBER: US/09/513,999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 9342
LENGTH: 271
TYPE: DNA
ORGANISM: Homo sapiens
US-09-513-999C-9342

Query Match 1.2%; Score 18; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 TCCTCACCACCAAGCCTCA 327
|||
Db 144 TCCTCACCACCAAGCCTCA 161

RESULT 13
US-09-641-638-549
Sequence 549, Application US/09641638
Patent No. 6432648
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
FILE REFERENCE: GENSET.051CP1
CURRENT APPLICATION NUMBER: US/09/641,638
CURRENT FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm

```
; SEQ ID NO 549
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 416
; OTHER INFORMATION: 12-63-402 : polymorphic base A or G
; NAME/KEY: misc_binding
; LOCATION: 396..415
; OTHER INFORMATION: 12-63-402.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 417..436
; OTHER INFORMATION: 12-63-402.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 15..35
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 445..464
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 404..428
; OTHER INFORMATION: 12-63-402 potential probe
; NAME/KEY: misc_feature
; LOCATION: 13..14
; OTHER INFORMATION: n=a, g, c or t
; US-09-641-638-549
```

```
Query Match 1.2%; Score 18; DB 3; Length 472;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 767 GAAATGAGCTGAGATT 784
Db 199 GAAATGAGCTGAGATT 216
```

```
RESULT 14
US-09-641-638-550
; Sequence 550, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueterec, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 550
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88
; OTHER INFORMATION: 12-63-74 : polymorphic base A or G
; NAME/KEY: misc_binding
; LOCATION: 68..87
; OTHER INFORMATION: 12-63-74.misl, potential
; NAME/KEY: misc_binding
```

```
; LOCATION: 89..108
; OTHER INFORMATION: 12-63-74.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 15..35
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 445..464
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 76..100
; OTHER INFORMATION: 12-63-74 potential probe
; NAME/KEY: misc_feature
; LOCATION: 13..14
; OTHER INFORMATION: n=a, g, c or t
; US-09-641-638-550
```

```
Query Match 1.2%; Score 18; DB 3; Length 472;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 767 GAAATGAGCTGAGATT 784
Db 199 GAAATGAGCTGAGATT 216
```

```
RESULT 15
US-10-170-097-549
; Sequence 549, Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueterec, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GEN-T14XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 549
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 416
; OTHER INFORMATION: 12-63-402 : polymorphic base A or G
; NAME/KEY: misc_binding
; LOCATION: 396..415
; OTHER INFORMATION: 12-63-402.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 417..436
; OTHER INFORMATION: 12-63-402.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 15..35
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
```

```
/ LOCATION: 445..464
/ OTHER INFORMATION: downstream amplification primer, complement
/ FEATURE:
/ NAME/KEY: misc.binding
/ LOCATION: 404..428
/ OTHER INFORMATION: 12-63-402 potential probe
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: 13..14
/ OTHER INFORMATION: n=a, g, c or t
US-10-170-097-549
```

```
Query Match 1.2%; Score 18; DB 4; Length 472;
Best Local Similarity 100.0%; Pred.No. 79;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 767 GAAATGAGGCTGAGATT 784
    |||||
Db 199 GAAATGAGGCTGAGATT 216
```

```
Search completed: June 8, 2005, 14:43:08
Job time : 260.698 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 11:26:41 ; Search time 950.618 Seconds
(without alignments)
9996.820 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533

Sequence: 1 agcttcacgcagaatttaa.....tgattctcgagracaaagta 1533

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 6046767 seqs, 3099530249 residues

Word size : 0

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Listing first 45 summaries

Database : Published Applications NA.*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14:	/cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16:	/cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17:	/cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18:	/cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19:	/cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20:	/cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21:	/cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22:	/cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq.*
23:	/cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24:	/cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1533	100.0	1533	14	US-10-025-003-11
2	1533	100.0	1533	14	US-10-718-952-11
3	1482	96.7	1533	14	US-10-025-003-13
4	1482	96.7	1533	14	US-10-025-003-15
5	1482	96.7	1533	14	US-10-718-952-13
6	1482	96.7	1533	19	US-10-718-952-15
7	986	64.3	2582	18	US-10-424-599-12021

8	310	10.2	634	18	US-10-424-599-22663	Sequence 22663, A
9	185	12.1	844	18	US-10-424-599-70162	Sequence 70162, A
10	146	9.5	661	18	US-10-424-599-116597	Sequence 116597, A
11	146	9.5	1533	14	US-10-025-003-5	Sequence 5, Appl1
12	146	9.5	1533	14	US-10-025-003-9	Sequence 9, Appl1
13	146	9.5	1533	14	US-10-718-952-5	Sequence 5, Appl1
14	146	9.5	1533	19	US-10-718-952-9	Sequence 9, Appl1
15	146	9.5	1760	14	US-10-025-003-1	Sequence 1, Appl1
16	146	9.5	1760	14	US-10-718-952-1	Sequence 1, Appl1
17	146	9.5	1989	18	US-10-424-599-70167	Sequence 70167, A
18	121	7.9	594	18	US-10-424-599-70166	Sequence 70166, A
19	44	2.9	2018	18	US-10-424-599-12022	Sequence 12022, A
20	35	2.3	989	18	US-10-424-599-12019	Sequence 12019, A
21	35	2.3	2582	18	US-10-424-599-12021	Sequence 12021, A
22	30	2.0	535	18	US-10-424-599-78882	Sequence 78882, A
23	28	1.8	355	18	US-10-424-599-94906	Sequence 94906, A
24	27	1.8	252	18	US-10-424-599-12017	Sequence 12017, A
25	26	1.7	377	11	US-09-733-6277A-2381	Sequence 2381, Ap
26	26	1.7	532	19	US-10-021-323-16419	Sequence 16419, A
27	26	1.7	533	19	US-10-021-323-16951	Sequence 16951, A
28	26	1.7	543	19	US-10-021-323-16651	Sequence 16651, A
29	26	1.7	566	19	US-10-021-323-12655	Sequence 12655, A
30	26	1.7	618	19	US-10-021-323-17771	Sequence 17771, A
31	26	1.7	1673	20	US-10-425-115-47180	Sequence 47180, A
32	26	1.7	1941	20	US-10-739-930-2091	Sequence 2091, Ap
33	24	1.6	451	18	US-10-424-599-137655	Sequence 137655, A
34	24	1.6	1924	14	US-10-767-701-14714	Sequence 14714, A
35	23	1.5	35	19	US-10-025-003-3	Sequence 3, Appl1
36	23	1.5	35	19	US-10-718-952-3	Sequence 3, Appl1
37	23	1.5	260	18	US-10-424-599-36573	Sequence 36573, A
38	23	1.5	290	19	US-10-437-963-16182	Sequence 16182, A
39	23	1.5	319	18	US-10-424-599-106050	Sequence 106050, A
40	23	1.5	345	11	US-09-732-6277A-3508	Sequence 3508, Ap
41	23	1.5	452	11	US-09-732-6277A-3572	Sequence 3572, Ap
42	23	1.5	557	19	US-10-021-323-12840	Sequence 12840, A
43	23	1.5	601	19	US-10-021-323-15240	Sequence 15240, A
44	23	1.5	605	19	US-10-021-323-15202	Sequence 15202, A
45	23	1.5	1898	18	US-10-425-114-7081	Sequence 7081, Ap

ALIGNMENTS

RESULT 1
US-10-025-003-11
; Sequence 11, Application US/10025003
; Publication No. US20030074685A1
GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025, 003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-11
Query Match 100.0%; Score 1533; DB 14; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ATGTCATCGAGAATTTTAAGGTAGAGAGTCTCTAATGTAGATACCGAGACTGAGATT 60
DB 1 ATGTCATCGAGAATTTTAAGGTAGAGAGTCTCTAATGTAGATACCGAGACTGAGATT 60
OY 61 CAGTCGGGTATACACTACGAAACACCGAACTTTGTTACAGAGAACAGAAATGGCACCTT 120
DB 61 CAGTCGGGTATACACTACGAAACACCGAACTTTGTTACAGAGAACAGAAATGGCACCTT 120
OY 121 CAGTGGATTTGTCAAAACCCAAATCCGTCATACCTAATTTAAACCAACACCCATGTTCCA 180
DB 121 CAGTGGATTTGTCAAAACCCAAATCCGTCATACCTAATTTAAACCAACACCCATGTTCCA 180
OY 181 AAATTGGGGGTATAGCTTGTGGGTGGGGTGAACCAACGAGCTTACCTCAACCGTGGT 240
DB 181 AAATTGGGGGTATAGCTTGTGGGTGGGGTGAACCAACGAGCTTACCTCAACCGTGGT 240
OY 241 GTTATTTGCTAACAGAGAGACATTTTATGTGGCTAACAGAGACAGATTAAACAGCCTAAT 300
DB 241 GTTATTTGCTAACAGAGAGACATTTTATGTGGCTAACAGAGACAGATTAAACAGCCTAAT 300
OY 301 TACTTTGGCTCCCTCAACCAAGCTCAGGATTTGAGTTGATCTTCCAGAGGAGAGAA 360
DB 301 TACTTTGGCTCCCTCAACCAAGCTCAGGATTTGAGTTGATCTTCCAGAGGAGAGAA 360
OY 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCATGTTAATCTGACGACATTTGTTGGG 420
DB 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCATGTTAATCTGACGACATTTGTTGGG 420
OY 421 GGATGGGATATACGACAACTGAACTGCTGATGCCATGGCCAGGGCAAAAGTGTTTGAC 480
DB 421 GGATGGGATATACGACAACTGAACTGCTGATGCCATGGCCAGGGCAAAAGTGTTTGAC 480
OY 481 ATCGATTTTGCAGAGAGATTGAGGCTTACATGGAATTCATGGTTCACTCCCGGAAATC 540
DB 481 ATCGATTTTGCAGAGAGATTGAGGCTTACATGGAATTCATGGTTCACTCCCGGAAATC 540
OY 541 TACGACCCCGAATTTCTATGCTGCAACCAAGAGAGCGTCCAAACACGTGATTAAGAGGC 600
DB 541 TACGACCCCGAATTTCTATGCTGCAACCAAGAGAGCGTCCAAACACGTGATTAAGAGGC 600
OY 601 ACAAAAGCAAGACAGTTCAGCAAAATCATCAAAAGACATCAAGGCGTTTAAAGAGCCACC 660
DB 601 ACAAAAGCAAGACAGTTCAGCAAAATCATCAAAAGACATCAAGGCGTTTAAAGAGCCACC 660
OY 661 AAAGTGAACAAGTGTGTCCTGCTGAGTGCACACAGAGAGGTATGCAATTTGTT 720
DB 661 AAAGTGAACAAGTGTGTCCTGCTGAGTGCACACAGAGAGGTATGCAATTTGTT 720
OY 721 GTAGGCTTATGACACATGAGAACTCTTGGCTGTGTGAGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTATGACACATGAGAACTCTTGGCTGTGTGAGACAGAAATGAGGCTGAG 780
OY 781 ATTTTCTCTTCCACTTGTATGCTGCTGTGTGTGAGAAAATTTCTTTCAATTAAT 840
DB 781 ATTTTCTCTTCCACTTGTATGCTGCTGTGTGTGAGAAAATTTCTTTCAATTAAT 840
OY 841 GGAAGCCCTCAGAACCTTTGTACAGAGGCTGATTTGATCTTGCCATTCGCGAGAGAACT 900
DB 841 GGAAGCCCTCAGAACCTTTGTACAGAGGCTGATTTGATCTTGCCATTCGCGAGAGAACT 900
OY 901 TTGATTTGTGAGATGACTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGTTGAT 960
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGAGACCAAAATGAAATCTGTGTTGTTGAT 960
OY 961 TTTCTTTGGGGGCTGGTATCAAGCAACATCTAATGTTAGTTAACACATCTGGGAAAC 1020
DB 961 TTTCTTTGGGGGCTGGTATCAAGCAACATCTAATGTTAGTTAACACATCTGGGAAAC 1020
OY 1021 AATGATNGATGAAATCTCGGCTCCAAACCTTCCGCTCCAAAGGAAATCTTCAAGAAC 1080
DB 1021 AATGATNGATGAAATCTCGGCTCCAAACCTTCCGCTCCAAAGGAAATCTTCAAGAAC 1080
OY 1081 AAGCTTGTGAGAGATGTGTCAACAGCAATGCAATCTCTATGAGCTGTGTGAACATCCC 1140
```

```
DB 1081 AACGTTGTGACGATATGTCAACAGCAATGCATCTCTATGAGCCCTGTGAAATCCC 1140
OY 1141 GACCATGTTGTTATTAAGATATGCTTACCGTAAAGGGGATAGCAAGAGCCATGGAT 1200
DB 1141 GACCATGTTGTTATTAAGATATGCTTACCGTAAAGGGGATAGCAAGAGCCATGGAT 1200
OY 1201 GAGTACACTTCAGAGATATTCATGAGGTGAAGAACACCATTTGTTTTCACAAACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTCATGAGGTGAAGAACACCATTTGTTTTCACAAACATGT 1260
OY 1261 GAGGATTCCTTTTATGCTGCTCTATTTATCTTGGACTTGTCTTCTGAGCTGAGC 1320
DB 1261 GAGGATTCCTTTTATGCTGCTCTATTTATCTTGGACTTGTCTTCTGAGCTGAGC 1320
OY 1321 ACTAAATCCAGTTTAAAGCTGAATAAGAGGAAATTCACCTCATCCACCACTTGGT 1380
DB 1321 ACTAAATCCAGTTTAAAGCTGAATAAGAGGAAATTCACCTCATCCACCACTTGGT 1380
OY 1381 ACCATTTCTCAGCTATCTGACCAAGGCTCTGTTCCACCGGGTACACAGTGTGAT 1440
DB 1381 ACCATTTCTCAGCTATCTGACCAAGGCTCTGTTCCACCGGGTACACAGTGTGAT 1440
OY 1441 GCATTGTCAAAAGCAGCTGCAATGCTGAAACATTAATGAGGCTTGTGTTGATTTGCC 1500
DB 1441 GCATTGTCAAAAGCAGCTGCAATGCTGAAACATTAATGAGGCTTGTGTTGATTTGCC 1500
OY 1501 CCAGAGAAATTAATGATTTCTGAGTACAAAGTGA 1533
DB 1501 CCAGAGAAATTAATGATTTCTGAGTACAAAGTGA 1533
```

RESULT 2

US-10-718-952-11

Sequence 11, Application US/10718952

Publication No. US20040128713A1

GENERAL INFORMATION:

APPLICANT: Hitz, William

APPLICANT: Sebastian, Scott

APPLICANT: Grace, John

APPLICANT: Street, Leon

TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE

FILE REFERENCE: BB-1077-C

CURRENT APPLICATION NUMBER: US/10/718,952

PRIOR FILING DATE: 2003-11-21

PRIOR APPLICATION NUMBER: 08/835,751

PRIOR FILING DATE: APRIL 8, 1997

PRIOR APPLICATION NUMBER: PCT/US98/06822

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Microsoft Office 97

SEQ ID NO 11

LENGTH: 1533

TYPE: DNA

ORGANISM: Glycine max

US-10-718-952-11

Query Match 100.0%; Score 1533; DB 19; Length 1533;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ATGTCATCGAGAATTTTAAGGTAGAGAGTCTCTAATGTAGATACCGAGACTGAGATT 60
DB 1 ATGTCATCGAGAATTTTAAGGTAGAGAGTCTCTAATGTAGATACCGAGACTGAGATT 60
OY 61 CAGTCGGGTATACACTACGAAACACCGAACTTTGTTACAGAGAACAGAAATGGCACCTT 120
DB 61 CAGTCGGGTATACACTACGAAACACCGAACTTTGTTACAGAGAACAGAAATGGCACCTT 120
OY 1201 GAGTACACTTCAGAGATATTCATGAGGTGAAGAACACCATTTGTTTTCACAAACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTCATGAGGTGAAGAACACCATTTGTTTTCACAAACATGT 1260
OY 1261 GAGGATTCCTTTTATGCTGCTCTATTTATCTTGGACTTGTCTTCTGAGCTGAGC 1320
DB 1261 GAGGATTCCTTTTATGCTGCTCTATTTATCTTGGACTTGTCTTCTGAGCTGAGC 1320
OY 1321 ACTAAATCCAGTTTAAAGCTGAATAAGAGGAAATTCACCTCATCCACCACTTGGT 1380
DB 1321 ACTAAATCCAGTTTAAAGCTGAATAAGAGGAAATTCACCTCATCCACCACTTGGT 1380
OY 1381 ACCATTTCTCAGCTATCTGACCAAGGCTCTGTTCCACCGGGTACACAGTGTGAT 1440
DB 1381 ACCATTTCTCAGCTATCTGACCAAGGCTCTGTTCCACCGGGTACACAGTGTGAT 1440
OY 1441 GCATTGTCAAAAGCAGCTGCAATGCTGAAACATTAATGAGGCTTGTGTTGATTTGCC 1500
DB 1441 GCATTGTCAAAAGCAGCTGCAATGCTGAAACATTAATGAGGCTTGTGTTGATTTGCC 1500
OY 1501 CCAGAGAAATTAATGATTTCTGAGTACAAAGTGA 1533
DB 1501 CCAGAGAAATTAATGATTTCTGAGTACAAAGTGA 1533
```


181 AATTTGGGGGTGATGCTTGTGGGTGGGGAACAACGGCTTACCCCTACCGGTGT 240
 181 AATTTGGGGGTGATGCTTGTGGGTGGGGAACAACGGCTTACCCCTACCGGTGT 240
 241 GTTATTGCTAACAGAGAGCATTTTCAATGGGCTTACAAAGAGCAAGATTTCAACAGCCAT 300
 241 GTTATTGCTAACAGAGAGCATTTTCAATGGGCTTACAAAGAGCAAGATTTCAACAGCCAT 300
 301 TACTTTGGCTCCCTCAACCGCTCAGTATTCAGATTCGATTCCTTCCAGGAGAGAA 360
 301 TACTTTGGCTCCCTCAACCGCTCAGTATTCAGATTCGATTCCTTCCAGGAGAGAA 360
 361 ATCTATGCCCCCATTTAAAGAGTCTGCTTCCAAATGTTAACTCTCAACGATTTGTTGGG 420
 361 ATCTATGCCCCCATTTAAAGAGTCTGCTTCCAAATGTTAACTCTCAACGATTTGTTGGG 420
 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCTGATGCTGATGCTGATGCTGATGCT 480
 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCTGATGCTGATGCTGATGCTGATGCT 480
 481 ATGATTTTGCAAGAGCATTTGAGGCTTACATGGAATCCATGTTCCACTCCCGAATC 540
 481 ATGATTTTGCAAGAGCATTTGAGGCTTACATGGAATCCATGTTCCACTCCCGAATC 540
 541 TACGACCCCGATTTTCAATGCTGCTCAACCAAGAGAGGCTGCTCAACCGATTTAAGGCG 600
 541 TACGACCCCGATTTTCAATGCTGCTCAACCAAGAGAGGCTGCTCAACCGATTTAAGGCG 600
 601 ACAAGCAAGAGAGCATTTGCAAAATCATCAAGAGCATCAAGGCGTTTAAAGAGGACCC 660
 601 ACAAGCAAGAGAGCATTTGCAAAATCATCAAGAGCATCAAGGCGTTTAAAGAGGACCC 660
 661 AAAAGTGAACAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 661 AAAAGTGAACAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 721 GTTATGCTTAAATCAACATGAGAGATCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 721 GTTATGCTTAAATCAACATGAGAGATCTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 781 ATTTCTCTTCCACCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 781 ATTTCTCTTCCACCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 841 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTCCATCGGAGGAACT 900
 841 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTCCATCGGAGGAACT 900
 901 TTGATTTGTTGAGATGATCTTCAAGAGTGTGCTGATGCTGATGCTGATGCTGATGCT 960
 901 TTGATTTGTTGAGATGATCTTCAAGAGTGTGCTGATGCTGATGCTGATGCTGATGCT 960
 961 TTTCTTGTGGGGCTGATGATCAAGCAACCTATGTTAGTTAAACAACCTTGGGAAAC 1020
 961 TTTCTTGTGGGGCTGATGATCAAGCAACCTATGTTAGTTAAACAACCTTGGGAAAC 1020
 1021 AATGATGATGATGATCTTCCGCTCCTCAAACTTCCGCTCCTCAAACTTCCGCTCCTCA 1080
 1021 AATGATGATGATGATCTTCCGCTCCTCAAACTTCCGCTCCTCAAACTTCCGCTCCTCA 1080
 1081 AAGCTTTGATGATGATGCTTCAAGAGTGTGCTGATGCTGATGCTGATGCTGATGCT 1140
 1081 AAGCTTTGATGATGATGCTTCAAGAGTGTGCTGATGCTGATGCTGATGCTGATGCT 1140
 1141 GACCATGTTGTTTATTTAAGTATGCTTACGTTAGGAGATGCAAGAGGCTGATGAT 1200
 1141 GACCATGTTGTTTATTTAAGTATGCTTACGTTAGGAGATGCAAGAGGCTGATGAT 1200
 1201 GAGTACATTTCAAGATATTCATGAGGTGAGAAACCACTTTGTTTGGCAACATGT 1260
 1201 GAGTACATTTCAAGATATTCATGAGGTGAGAAACCACTTTGTTTGGCAACATGT 1260

1261 GAGGATTCCTTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 1261 GAGGATTCCTTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 1321 ACTAGATTCAGATTTTAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
 1321 ACTAGATTCAGATTTTAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
 1381 ACCATTCAGATTTTAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
 1381 ACCATTCAGATTTTAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
 1441 GATTTGCTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
 1441 GATTTGCTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
 1501 CCAGAGATTAACATGATTTCTCGATCAAGTGA 1533
 1501 CCAGAGATTAACATGATTTCTCGATCAAGTGA 1533

RESULT 3
 US-10-025-003-13
 ; Sequence 13, Application US/10025003
 ; Publication No. US20030074685A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hitz, William
 ; APPLICANT: Sebastian, Scott
 ; APPLICANT: Grace, John
 ; APPLICANT: Streif, Leon
 ; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
 ; FILE REFERENCE: BB-1077-C
 ; CURRENT APPLICATION NUMBER: US/10/025, 003
 ; CURRENT FILING DATE: 2002-05-07
 ; PRIOR APPLICATION NUMBER: 08/835, 751
 ; PRIOR FILING DATE: APRIL 9, 1997
 ; PRIOR APPLICATION NUMBER: PCT/US98/06822
 ; PRIOR FILING DATE: APRIL 7, 1998
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 13
 ; LENGTH: 1533
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; US-10-025-003-13

Query Match 96.7%; Score 1482; DB 14; Length 1533;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ANTTTCATGAGATTTTAAAGTGAAGTCTTAAATGTAAGTACACCGAGCTGAGATT 60
 1 ANTTTCATGAGATTTTAAAGTGAAGTCTTAAATGTAAGTACACCGAGCTGAGATT 60
 61 CAGTCCGTGTAACACTAGCAACCGAATCTTGTTCAGGAGACGAAATGACCTAT 120
 61 CAGTCCGTGTAACACTAGCAACCGAATCTTGTTCAGGAGACGAAATGACCTAT 120
 121 CAGTGAATGTAACACCGAATCCGTCATCTAACCAATTTAAACCAACCGAATGTTCA 180
 121 CAGTGAATGTAACACCGAATCCGTCATCTAACCAATTTAAACCAACCGAATGTTCA 180
 181 AAATGGGGGATGATGCTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTT 240
 181 AAATGGGGGATGATGCTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTT 240
 241 GTTATTGCTAACAGAGAGCATTTTCAATGGGCTTACAAAGAGCAAGATTTCAACAGCCAT 300
 241 GTTATTGCTAACAGAGAGCATTTTCAATGGGCTTACAAAGAGCAAGATTTCAACAGCCAT 300
 301 TACTTTGGCTCCCTCAACCGCTCAGTATTCAGATTCGATTCCTTCCAGGAGAGAA 360
 301 TACTTTGGCTCCCTCAACCGCTCAGTATTCAGATTCGATTCCTTCCAGGAGAGAA 360

Db 301 TACTTTGGCTCCCTCAACCCAAAGCTCAGCTATTGAGTTGATCTCTCCAGGGAGAGAA 360
Qy ATCTATGCCCATTTGAAGAGTGTGCTTCAATGGTTAATCCGAGACATGTGTGTTGGG 420
Db 361 ATCTATGCCCATTTGAAGAGTGTGCTTCAATGGTTAATCCGAGACATGTGTGTTGGG 420
Qy 421 GGATGGGATATCAGCAACATGAACTGTGCTGATGTCATGCGCAAGGGATGTTTGAC 480
Db 421 GGATGGGATATCAGCAACATGAACTGTGCTGATGTCATGCGCAAGGGATGTTTGAC 480
Qy 481 ATCGATTTGCGAGACAGTTGAGGCTTTCATGAGTAATCCATGTTTCCACTCCCCGGAATC 540
Db 481 ATCGATTTGCGAGACAGTTGAGGCTTTCATGAGTAATCCATGTTTCCACTCCCCGGAATC 540
Qy 541 TACGACCCGGATTTTCAATGCTGTGCTCAACCAAGAGAGCGGTGCAACACGTTAATGAAGGC 600
Db 541 TACGACCCGGATTTTCAATGCTGTGCTCAACCAAGAGAGCGGTGCAACACGTTAATGAAGGC 600
Qy 601 ACAAGCAAGACAGATTCAGCAAAATCATCAAAAGCATCAAGCGCTTTAAGAGACCAAC 660
Db 601 ACAAGCAAGACAGATTCAGCAAAATCATCAAAAGCATCAAGCGCTTTAAGAGACCAAC 660
Qy 661 AAAGTGACAGAGTGTGTTCTCTGTGAGTGTGCAACACAGAGATATGCAATTTGGTT 720
Db 661 AAAGTGACAGAGTGTGTTCTCTGTGAGTGTGCAACACAGAGATATGCAATTTGGTT 720
Qy 721 GTAGGCTTTAATGACACCATGAGAAATCTCTGGGCTGTGTGAGACAGAAATAGGCTGAG 780
Db 721 GTAGGCTTTAATGACACCATGAGAAATCTCTGGGCTGTGTGAGACAGAAATAGGCTGAG 780
Qy 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGATGGAATAATGTTCTTTCAATTAAT 840
Db 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGATGGAATAATGTTCTTTCAATTAAT 840
Qy 841 GGAAGCCTCAGAAACCTTTTGACAGAGGCTGATTTGATCTTCCATCGCAGAGAACT 900
Db 841 GGAAGCCTCAGAAACCTTTTGACAGAGGCTGATTTGATCTTCCATCGCAGAGAACT 900
Qy 901 TTGATTTGGTGAGATGACTTCAAGAGTGTGTCAGACCAAAATGAATCTGTGTTGGTAT 960
Db 901 TTGATTTGGTGAGATGACTTCAAGAGTGTGTCAGACCAAAATGAATCTGTGTTGGTAT 960
Qy 961 TTTCTTGGGGGCTGTGATCAAGCCACATCTATAGTTAGTTACAAACCATCTGGGAAAC 1020
Db 961 TTTCTTGGGGGCTGTGATCAAGCCACATCTATAGTTAGTTACAAACCATCTGGGAAAC 1020
Qy 1021 AATGATGATGATATCTCTGCGCTCCAAACCTTCCGCTCAAGAAATCTTCCAGAGC 1080
Db 1021 AATGATGATGATATCTCTGCGCTCCAAACCTTCCGCTCAAGAAATCTTCCAGAGC 1080
Qy 1081 AAGGTTTGAAGATATGATCAACAGCAATGCCATCTCTATGAGCGCTGTGAAACATCCC 1140
Db 1081 AAGGTTTGAAGATATGATCAACAGCAATGCCATCTCTATGAGCGCTGTGAAACATCCC 1140
Qy 1141 GACCAATGTTGTTAATTAAGATATGCTTACGTAAGGAGATAGCAAGAGCCATGAT 1200
Db 1141 GACCAATGTTGTTAATTAAGATATGCTTACGTAAGGAGATAGCAAGAGCCATGAT 1200
Qy 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAAGAACATGTTTGTGCAACACATGT 1260
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAAGAACATGTTTGTGCAACACATGT 1260
Qy 1261 GAGGATTCCTTTTGGCTGCTCTATATCTTGGATCTTGCTTGTGAGAGCTGAGAC 1320
Db 1261 GAGGATTCCTTTTGGCTGCTCTATATCTTGGATCTTGCTTGTGAGAGCTGAGAC 1320
Qy 1321 ACTAGATATCAGTTTAAAGCTGAAAATGAGGAAAATTCACATCTTCCAGCAGTTGCT 1380
Db 1321 ACTAGATATCAGTTTAAAGCTGAAAATGAGGAAAATTCACATCTTCCAGCAGTTGCT 1380
Qy 1381 ACCATTTCTAGCTATATGACCAAGGCTCTCTGTGTTCCACCGGGTACACCAAGTGTGAAT 1440
Db 1381 ACCATTTCTAGCTATATGACCAAGGCTCTCTGTGTTCCACCGGGTACACCAAGTGTGAAT 1440

Qy 1441 GCATTGTCAAAAGCAGCGTCAATGCTGGAATAATAGAGGCTTGTGTGATTTGGCC 1500
Db 1441 GCATTGTCAAAAGCAGCGTCAATGCTGGAATAATAGAGGCTTGTGTGATTTGGCC 1500
Qy 1501 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1533
Db 1501 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1533

RESULT 4
US-10-025-003-15
; Sequence 15, Application US/10025003
; Publication No. US2003074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streil, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BR-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR FILING DATE: 08/835,751
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-15

Query Match 96.7%; Score 1482; DB 14; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGTTTCATCGAGATTTTAAAGTATAGAGTCTCTATATGGAATTAACGAGATGAGATT 60
Db 1 ATGTTTCATCGAGATTTTAAAGTATAGAGTCTCTATATGGAATTAACGAGATGAGATT 60
Qy 61 CAGTCGCTGTAACATTAAGAAACACCGAATCTGTTCAAGAAACAGAAATGSCACTAT 120
Db 61 CAGTCGCTGTAACATTAAGAAACACCGAATCTGTTCAAGAAACAGAAATGSCACTAT 120
Qy 121 CAGTGAATTTGCAACCAAAATCGGTCAATCAATTTAAACCAACACCATGTTCCA 180
Db 121 CAGTGAATTTGCAACCAAAATCGGTCAATCAATTTAAACCAACACCATGTTCCA 180
Qy 181 AAATTGGGGGTATGCTTGTGGTGGGTGGAACCAACGGTCTAACCCCTGACCGGTGT 240
Db 181 AAATTGGGGGTATGCTTGTGGTGGGTGGAACCAACGGTCTAACCCCTGACCGGTGT 240
Qy 241 GTTATTTGCTAACAGAGAGACATTTCAATGGGCTACAAAGAACAAATTTCAACCAAT 300
Db 241 GTTATTTGCTAACAGAGAGAGACATTTCAATGGGCTACAAAGAACAAATTTCAACCAAT 300
Qy 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATCTTCCAGGAGAGAA 360
Db 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATCTTCCAGGAGAGAA 360
Qy 361 ATCTATGCCCATTTGAAGAGTGTGCTTCAATGGTTAATCCGAGACATGTGTGTTGGG 420
Db 361 ATCTATGCCCATTTGAAGAGTGTGCTTCAATGGTTAATCCGAGACATGTGTGTTGGG 420
Qy 421 GGATGGGATATCAGCAACATGAACTGTGCTGATGTCATGCGCAAGGGATGTTTGAC 480
Db 421 GGATGGGATATCAGCAACATGAACTGTGCTGATGTCATGCGCAAGGGATGTTTGAC 480
Qy 481 ATCGATTTGCGAGACAGTTGAGGCTTTCATGAGTAATCCATGTTTCCACTCCCCGGAATC 540

```

481 ATGATTTGAGAGAGAGGAGGCTTACATGGAATCCATGGTTCCACTCCCGGAATC 540
541 TAGAGCCCGGATTTCTTGTCTGCCAACAGAGAGGCTGCCAACACGTGATTAAGGCG 600
541 TAGAGCCCGGATTTCTTGTCTGCCAACAGAGAGGCTGCCAACACGTGATTAAGGCG 600
601 ACAAGCAGAGAGAGGAGGCTTACATGGAATCCATGGGCTTTAAGAGAGGCGAC 660
601 ACAAGCAGAGAGGAGGCTTACATGGAATCCATGGGCTTTAAGAGAGGCGAC 660
661 AAAGTGAAGAGGAGGCTTCTGTGAGCTGCCAACAGAGAGGATTAAGCATTTGGT 720
661 AAAGTGAAGAGGAGGCTTCTGTGAGCTGCCAACAGAGAGGATTAAGCATTTGGT 720
721 GTAGGCTTTAATGACACATGAGAAATCTTGGCTGTGAGCAGAAATGAGGCTGAG 780
721 GTAGGCTTTAATGACACATGAGAAATCTTGGCTGTGAGCAGAAATGAGGCTGAG 780
781 ATTTCTCTTCCACCTTGTATGCGAATGCTGTGATGAGAAATGTTCTTCAATTAAT 840
781 ATTTCTCTTCCACCTTGTATGCGAATGCTGTGATGAGAAATGTTCTTCAATTAAT 840
841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTGCATCGGAGAGCACT 900
841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTGCATCGGAGAGCACT 900
901 TTGATTTGAGAGAGTACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTTGAT 960
901 TTGATTTGAGAGAGTACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTTGAT 960
961 TTTCTTGTGGGGCTGTATCAAGCCCAATCTATGTTAGTACAAACCATCTGGGAAC 1020
961 TTTCTTGTGGGGCTGTATCAAGCCCAATCTATGTTAGTACAAACCATCTGGGAAC 1020
1021 AATGATGATGATCTCTCGGCTCCAAACCTTCCGCTCCAAAGGAAATCTTCAAGAG 1080
1021 AATGATGATGATCTCTCGGCTCCAAACCTTCCGCTCCAAAGGAAATCTTCAAGAG 1080
1081 AAGCTTTGAGAGATTTGCTCAACAGCAATGCCATCTCTATGAGCTGTGAAATCCC 1140
1081 AAGCTTTGAGAGATTTGCTCAACAGCAATGCCATCTCTATGAGCTGTGAAATCCC 1140
1141 GACCATTTGTTTATTAAGTATGCTTACGTAGGGGATGACAAAGAGCATGAT 1200
1141 GACCATTTGTTTATTAAGTATGCTTACGTAGGGGATGACAAAGAGCATGAT 1200
1201 GAGTACACTTCAGAGATTTTCAATGGGTGAAAGAACACATTTGTTGCAACACATGT 1260
1201 GAGTACACTTCAGAGATTTTCAATGGGTGAAAGAACACATTTGTTGCAACACATGT 1260
1261 GAGGATTCCTTTTACGCTCTATTAATCTTGGAATTTGCTTCTGTGAGCTGAGC 1320
1261 GAGGATTCCTTTTACGCTCTATTAATCTTGGAATTTGCTTCTGTGAGCTGAGC 1320
1321 ACTAAGATTCAGTTTAACTGAAAATGAGGAAAATTCATCTCAATCCACCATTTGCT 1380
1321 ACTAAGATTCAGTTTAACTGAAAATGAGGAAAATTCATCTCAATCCACCATTTGCT 1380
1381 ACCATTTCTAGCTATTTGACAAAGGCTCTGTGTTCCACCGGATGACAGTGTGAT 1440
1381 ACCATTTCTAGCTATTTGACAAAGGCTCTGTGTTCCACCGGATGACAGTGTGAT 1440
1441 GATTTGTCAAGAGGCTGCAATGCTGAAAACATATAGAGGCTTTGTGATTTGGCC 1500
1441 GATTTGTCAAGAGGCTGCAATGCTGAAAACATATAGAGGCTTTGTGATTTGGCC 1500
1501 CCAAGAAATTAATGATTTCTGAGTCAAGTGA 1533
1501 CCAAGAAATTAATGATTTCTGAGTCAAGTGA 1533

```

RESULT 5

```

US-10-718-952-13
; Sequence 13, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718, 952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-13

```

Query Match 96.7%; Score 1482; DB 19; Length 1533;

Best Local Similarity 99.9%; Pred. No. 0; Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

1 ATGTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
1 ATGTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
61 CAGTCGTGTACACTACGAAACCAACGAACTTGTTCAGAGAACAGAAATGAGCACTAT 120
61 CAGTCGTGTACACTACGAAACCAACGAACTTGTTCAGAGAACAGAAATGAGCACTAT 120
121 CAGTGGATTTGTAACCAACCCCAATCCGCTCACTACCAATTTAAACCAACCAATTTTCA 180
121 CAGTGGATTTGTAACCAACCCCAATCCGCTCACTACCAATTTAAACCAACCAATTTTCA 180
181 AAATTTGGGGGATGATCTTGTGGGTGGGATGAAACCAAGGCTTACCTGACCGGTGAT 240
181 AAATTTGGGGGATGATCTTGTGGGTGGGATGAAACCAAGGCTTACCTGACCGGTGAT 240
241 GTTATTTGCTTACAGAGAGACATTTCAATGGGCTTACAAAGACAAAGATTTCAACAGCAAT 300
241 GTTATTTGCTTACAGAGAGACATTTCAATGGGCTTACAAAGACAAAGATTTCAACAGCAAT 300
301 TACTTTGGCTCCCTCAACCAAGCTTCAAGCTTCAAGTGTGATCTTCCAGAGAGAGAA 360
301 TACTTTGGCTCCCTCAACCAAGCTTCAAGCTTCAAGTGTGATCTTCCAGAGAGAGAA 360
361 ATCTAATGCCCATTTAAGAGTGTGCTTCAATGGTTAATCTGAGAGCAATTTGTTGGG 420
361 ATCTAATGCCCATTTAAGAGTGTGCTTCAATGGTTAATCTGAGAGCAATTTGTTGGG 420
421 GGATGGGATTAAGCAATGAACTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 480
421 GGATGGGATTAAGCAATGAACTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 480
481 ATGATTTGACAGAGAGAGTGAAGGCTTCAATGAAATCCATGTTTCACTCCCGGAATC 540
481 ATGATTTGACAGAGAGAGTGAAGGCTTCAATGAAATCCATGTTTCACTCCCGGAATC 540
541 TAGAGCCCGGATTTATTTGCTGCAACAGAGAGGCTGCAACAGAGGCTTTAAGAGGCG 600
541 TAGAGCCCGGATTTATTTGCTGCAACAGAGAGGCTGCAACAGAGGCTTTAAGAGGCG 600
601 ACAAGCAGAGAGAGGCTTACATGGAATCCATGGGCTTTAAGAGAGGCGAC 660
601 ACAAGCAGAGAGAGGCTTACATGGAATCCATGGGCTTTAAGAGAGGCGAC 660

```

Qy 661 AAAGTGACAGAGTGTGTCTCTGTGAGTCCCAACAGAGAGTATGCAATTTGGTT 720
Db 661 AAAGTGACAGAGTGTGTCTCTGTGAGTCCCAACAGAGAGTATGCAATTTGGTT 720
Qy 721 GTAGGCTTTAATGACACCAATGAGAAATCTCTGCTGCTGTGACAGAAATAGGCTGAG 780
Db 721 GTAGGCTTTAATGACACCAATGAGAAATCTCTGCTGCTGTGACAGAAATAGGCTGAG 780
Qy 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGATGGAATAATGTTCTTTCAATTAAT 840
Db 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGATGGAATAATGTTCTTTCAATTAAT 840
Qy 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTCCATGCGAGAACCT 900
Db 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTCCATGCGAGAACCT 900
Qy 901 TTGATTTGTGAGAGTACCTTCAAGAGTGTGTGACACCAATAATGATGTTGTGAT 960
Db 901 TTGATTTGTGAGAGTACCTTCAAGAGTGTGTGACACCAATAATGATGTTGTGAT 960
Qy 961 TTTCTTTGGGGGCTGTATCAAGCCCAACATATAGTTAGTACCAACCATCTGGGAAAC 1020
Db 961 TTTCTTTGGGGGCTGTATCAAGCCCAACATATAGTTAGTACCAACCATCTGGGAAAC 1020
Qy 1021 AATGATGTATGATCTCTGAGCTCCCAACCTTCCGCTCCAGAGAAATCTCCAGAGC 1080
Db 1021 AATGATGTATGATCTCTGAGCTCCCAACCTTCCGCTCCAGAGAAATCTCCAGAGC 1080
Qy 1081 AACGTTGTGAGATATGCTCAACAGCAATGCCATCTCTTATAGAGCTGTGTAACATCCC 1140
Db 1081 AACGTTGTGAGATATGCTCAACAGCAATGCCATCTCTTATAGAGCTGTGTAACATCCC 1140
Qy 1141 GACCATGTTGTTATTAATGATGCTTACGTAAGGAGATAGCAAGAGACCATGAT 1200
Db 1141 GACCATGTTGTTATTAATGATGCTTACGTAAGGAGATAGCAAGAGACCATGAT 1200
Qy 1201 GAGTACACTTCAAGATATTCATGAGTGAAGAAACACCATTTGTTGCAACACATGT 1260
Db 1201 GAGTACACTTCAAGATATTCATGAGTGAAGAAACACCATTTGTTGCAACACATGT 1260
Qy 1261 GAGGATTCCTTTTGTGCTCTCTATTAATCTTGAAGTGTCTTCTGTGAGAGC 1320
Db 1261 GAGGATTCCTTTTGTGCTCTCTATTAATCTTGAAGTGTCTTCTGTGAGAGC 1320
Qy 1321 ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCACATCTCCACCAAGTGT 1380
Db 1321 ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCACATCTCCACCAAGTGT 1380
Qy 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTCTGTTCCACCGGATACCAAGTGTGAAT 1440
Db 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTCTGTTCCACCGGATACCAAGTGTGAAT 1440
Qy 1441 GATTTGTCAAAAGAGGTGCAATGCTGGAATAATGAGGCTGTGTGATTTGGCC 1500
Db 1441 GATTTGTCAAAAGAGGTGCAATGCTGGAATAATGAGGCTGTGTGATTTGGCC 1500
Qy 1501 CCAGAGATATACATGATTTCTGAGTACAGTGA 1533
Db 1501 CCAGAGATATACATGATTTCTGAGTACAGTGA 1533

RESULT 6
US-10-718-952-15
; Sequence 15, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C

Qy 1 ATGTTTCATGAGATTTTAAAGTATGAGTCTTAAATGTGAAATGACACCGAGACTGAGATT 60
Db 1 ATGTTTCATGAGATTTTAAAGTATGAGTCTTAAATGTGAAATGACACCGAGACTGAGATT 60
Qy 61 CAGTCCGTGTCACACTAGAGAAACACGAGCTGTTCACGAGAACAGGATGGACCTAT 120
Db 61 CAGTCCGTGTCACACTAGAGAAACACGAGCTGTTCACGAGAACAGGATGGACCTAT 120
Qy 61 CAGTCCGTGTCACACTAGAGAAACACGAGCTGTTCACGAGAACAGGATGGACCTAT 120
Db 61 CAGTCCGTGTCACACTAGAGAAACACGAGCTGTTCACGAGAACAGGATGGACCTAT 120
Qy 121 CAGTGGATTTGTCACCAACCAATCCGTCACTACCAATTTAAACCAACCCATGTTTCA 180
Db 121 CAGTGGATTTGTCACCAACCAATCCGTCACTACCAATTTAAACCAACCCATGTTTCA 180
Qy 181 AAATTTGGGGGTGATGCTTTGAGTTGGGGGTGAAACAAAGGCTCTACCTTCAAGGTTGT 240
Db 181 AAATTTGGGGGTGATGCTTTGAGTTGGGGGTGAAACAAAGGCTCTACCTTCAAGGTTGT 240
Qy 241 GTTATTTGTAACAGAGAGATTTTATGAGGCTTCAAGAGCAAGATTCAACAGCAAT 300
Db 241 GTTATTTGTAACAGAGAGATTTTATGAGGCTTCAAGAGCAAGATTCAACAGCAAT 300
Qy 301 TACTTTGGCTCCTTCAACCAAGCTTCAAGTATTCAGATTTGATCTTCCAGAGAGAGAA 360
Db 301 TACTTTGGCTCCTTCAACCAAGCTTCAAGTATTCAGATTTGATCTTCCAGAGAGAGAA 360
Qy 361 ATCTATGCCCCATTTAAGATGCTGTTCAATGTTAATCTGTACAGACATTTGTGTTGGG 420
Db 361 ATCTATGCCCCATTTAAGATGCTGTTCAATGTTAATCTGTACAGACATTTGTGTTGGG 420
Qy 421 GGAATGGATATCAGAAATGAACTGCTGATGCAATGCGCAAGGCAAGAGTGTGATG 480
Db 421 GGAATGGATATCAGAAATGAACTGCTGATGCAATGCGCAAGGCAAGAGTGTGATG 480
Qy 481 ATCGATTTGCAAGAGCAATGAGGCTTACATGAAATCATGATGTTTCCACTCCCGGAATC 540
Db 481 ATCGATTTGCAAGAGCAATGAGGCTTACATGAAATCATGATGTTTCCACTCCCGGAATC 540
Qy 541 TACGACCCGGATTTTATGCTGCTGCAACCAAGAGGCTGTGCAACAGATATTAAGGGC 600
Db 541 TACGACCCGGATTTTATGCTGCTGCAACCAAGAGGCTGTGCAACAGATATTAAGGGC 600
Qy 601 ACAAGCAAGAGCAAGTTCAGCAATCATCAAGACATCAAGAGGCTTTAAAGAGCCACC 660
Db 601 ACAAGCAAGAGCAAGTTCAGCAATCATCAAGACATCAAGAGGCTTTAAAGAGCCACC 660
Qy 661 AAAGTGACAGAGTGTGTCTCTGTGAGTCCCAACAGAGAGTATGCAATTTGGTT 720
Db 661 AAAGTGACAGAGTGTGTCTCTGTGAGTCCCAACAGAGAGTATGCAATTTGGTT 720
Qy 721 GTAGGCTTTAATGACACCAATGAGAAATCTCTGCTGCTGTGACAGAAATAGGCTGAG 780
Db 721 GTAGGCTTTAATGACACCAATGAGAAATCTCTGCTGCTGTGACAGAAATAGGCTGAG 780
Qy 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGATGGAATAATGTTCTTTCAATTAAT 840
Db 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGATGGAATAATGTTCTTTCAATTAAT 840

```
QY 841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTTGATTCGCAATCGGAGAAACACT 900
DB 841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTTGATTCGCAATCGGAGAAACACT 900
QY 901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTTGAT 960
DB 901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTTGAT 960
QY 961 TTTCTTTGGGGGCTGTGTATCAAGCCAACTATAGTTAGTTACCAACATCTGGGAAAC 1020
DB 961 TTTCTTTGGGGGCTGTGTATCAAGCCAACTATAGTTAGTTACCAACATCTGGGAAAC 1020
QY 1021 AATGATGTATGAATCTCTCGGCTCCAAACAACTTCGCTCCAAAGAAATCTCCAGAGC 1080
DB 1021 AATGATGTATGAATCTCTCGGCTCCAAACAACTTCGCTCCAAAGAAATCTCCAGAGC 1080
QY 1081 AAGCTTTGACAGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGAAACATCCC 1140
DB 1081 AAGCTTTGACAGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGAAACATCCC 1140
QY 1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGGGATAGCAAGAGCCATGAT 1200
DB 1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGGGATAGCAAGAGCCATGAT 1200
QY 1201 GAGTACACTTCAGAGATATTCAGAGGTGAGAAACACCATTTGTTGACACACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTCAGAGGTGAGAAACACCATTTGTTGACACACATGT 1260
QY 1261 GAGGATTTCCCTTTTAACTGCTCCTATTTATCTTGAATGATCTTCTGTGAGCTGAGC 1320
DB 1261 GAGGATTTCCCTTTTAACTGCTCCTATTTATCTTGAATGATCTTCTGTGAGCTGAGC 1320
QY 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACACTTCCATCCAGTTGCT 1380
DB 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACACTTCCATCCAGTTGCT 1380
QY 1381 ACCATTTCTAGCATATGTGACCAAGGCTCCTCTGTGTTCCACCGGTACACAGTGTGAT 1440
DB 1381 ACCATTTCTAGCATATGTGACCAAGGCTCCTCTGTGTTCCACCGGTACACAGTGTGAT 1440
QY 1441 GCATTGTCAAGACAGCGTCAATGCTGAAACATATGAGGCTGTGTGATTTGGCC 1500
DB 1441 GCATTGTCAAGACAGCGTCAATGCTGAAACATATGAGGCTGTGTGATTTGGCC 1500
QY 1501 CCAGAGATTAACATGATTTCTCGAGTACAACTGA 1533
DB 1501 CCAGAGATTAACATGATTTCTCGAGTACAACTGA 1533
```

RESULT 7

```
US-10-424-599-12021/c
; Sequence 12021, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12021
; LENGTH: 2582
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110862C.1
US-10-424-599-12021
```

```
Query Match 64.3%; Score 986; DB 18; Length 2582;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1136; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGTTTCATGAGAAATTTTAAAGTATGAGAGTCTTAAATGTGAATACACCGAGACTGAGATT 60
DB 2454 ATGTTTCATGAGAAATTTTAAAGTATGAGAGTCTTAAATGTGAATACACCGAGACTGAGATT 2395
QY 61 CAGTCCGTGTACACTACGAAACCAACCGAATCTTTCACGAGAACGGAATGGCACTAT 120
DB 2394 CAGTCCGTGTACACTACGAAACCAACCGAATCTTTCACGAGAACGGAATGGCACTAT 2335
QY 121 CAGTGGATTTGTAACCCCAATCCGCTCACTACCAATTTAAACCAACCCCACTGTCCA 180
DB 2334 CAGTGGATTTGTAACCCCAATCCGCTCACTACCAATTTAAACCAACCCCACTGTCCA 2275
QY 181 AATTTGGGGGTATGTTGTGGGTTGGGGTGGAAACAACGCTTACCTCAACCGGTGT 240
DB 2274 AATTTGGGGGTATGTTGTGGGTTGGGGTGGAAACAACGCTTACCTCAACCGGTGT 2215
QY 241 GTTATTTGCTAACAGAGAGCAATTTTCATAGGCTTACAAAGACAGATTCACAGCCAT 300
DB 2214 GTTATTTGCTAACAGAGAGCAATTTTCATAGGCTTACAAAGACAGATTCACAGCCAT 2155
QY 301 TACTTTGGCTCCCTACCCCAAGCTCAGTATTCAGATTTGATTCCTTCCAGGAGAGAA 360
DB 2154 TACTTTGGCTCCCTACCCCAAGCTCAGTATTCAGATTTGATTCCTTCCAGGAGAGAA 2095
QY 361 ATCTATGCCCCCATTAAGAGTCTGCTTCCATGTTTAACTCTGACGACATTTGTTGGG 420
DB 2094 ATCTATGCCCCCATTAAGAGTCTGCTTCCATGTTTAACTCTGACGACATTTGTTGGG 2035
QY 421 GATGGGATATACGAAACATGAACCTGTGTATGCCATGACGAGCAAGGATGTTTGA 480
DB 2034 GATGGGATATACGAAACATGAACCTGTGTATGCCATGACGAGCAAGGATGTTTGA 1975
QY 481 ATCGATTTGCAAGACAGTTGAGGCTTACATGAAATCCATGTTCCACTCCCGGAATC 540
DB 1974 ATCGATTTGCAAGACAGTTGAGGCTTACATGAAATCCATGTTCCACTCCCGGAATC 1915
QY 541 TACGACCCGGAATTTTCATTTGCTGCCAACCAAGAGAGCGGCCAACACGTGATTAAGG 600
DB 1914 TACGACCCGGAATTTTCATTTGCTGCCAACCAAGAGAGCGGCCAACACGTGATTAAGG 1855
QY 601 ACAAAAGCAAGACCAATTCAGCAAAATCATCAAGACATCAAGCGTTTAAAGAGCCACC 660
DB 1854 ACAAAAGCAAGACCAATTCAGCAAAATCATCAAGACATCAAGCGTTTAAAGAGCCACC 1795
QY 661 AAGGTGACAAAGTGTGTCTGTGACTGCAACACAGAGATTTGCAATTTGGTT 720
DB 1794 AAGGTGACAAAGTGTGTCTGTGACTGCAACACAGAGATTTGCAATTTGGTT 1735
QY 721 GTAGGCTTTAATGACACCATGAGAAATCTTGGGCTGTGAGACAGAAATGAGGCTGAG 780
DB 1734 GTAGGCTTTAATGACACCATGAGAAATCTTGGGCTGTGAGACAGAAATGAGGCTGAG 1675
QY 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGTGATGAAATGTTCTTTCAATTAAT 840
DB 1674 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGTGATGAAATGTTCTTTCAATTAAT 1615
QY 841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTTGATTCGCAATCGGAGAAACACT 900
DB 1614 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTTGATTCGCAATCGGAGAAACACT 1555
QY 901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTTGAT 960
DB 1554 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTTGAT 1495
QY 961 TTTCTTTGGGGGCTGTGTATCAAGCCAACTATAGTTAGTTACCAACATCTGGGAAAC 1020
DB 1494 TTTCTTTGGGGGCTGTGTATCAAGCCAACTATAGTTAGTTACCAACATCTGGGAAAC 1435
QY 1021 AATGATGTATGAATCTCTCGGCTCCAAACAACTTCGCTCCAAAGAAATCTCCAGAGC 1080
```

Db 1434 AATGATGATGAATCTCGGCTCCACAAACCTTCGCTCCAGGAATCTCCAGAGC 1375
QY 1081 AACGTTGTGAGATATGCTCAACAGCAATGCACTTCCTATAGAGCCTGGTGAACATCC 1139
Db 1374 AACGTTGTGAGATATGCTCAACAGCAATGCACTTCCTATAGAGCCTGGTGAACATCC 1316

RESULT 8
US-10-424-599-22663

; Sequence 22663, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 22663
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_120467C.1
US-10-424-599-22663

Query Match 20.2%; Score 310; DB 18; Length 634;
Best Local Similarity 100.0%; Pred. No. 8.3e-156;

Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 705 GTATAGCAATTTGGTTGATGAGCCTTATGACACCATGAGGAATCTCTGGCTGCTGGA 764
Db 4 GTATAGCAATTTGGTTGATGAGCCTTATGACACCATGAGGAATCTCTGGCTGCTGGA 63
QY 765 CAGGAATGAGGCTGATTTCTCCCTCCCTGATATGCACTTCCTGCTGATGAGGA 824
Db 64 CAGGAATGAGGCTGATTTCTCCCTCCCTGATATGCACTTCCTGCTGATGAGGA 123
QY 825 TGTTCCTTCATTAATGGAAGCCTCAGAACATTTTGTACAGAGGCTGATGATCTTGC 884
Db 124 TGTTCCTTCATTAATGGAAGCCTCAGAACATTTTGTACAGAGGCTGATGATCTTGC 183
QY 885 CATCGAGAGAACATTTGATGAGATGATGATCAAGAGTGTGACAGCAAAATGAA 944
Db 184 CATCGAGAGAACATTTGATGAGATGATGATCAAGAGTGTGACAGCAAAATGAA 243
QY 945 ATCTGTGTTGTTGATTTCTTGTGGGGGCTGATCAAGCCAACTATATGTTAGTTA 1004
Db 244 ATCTGTGTTGTTGATTTCTTGTGGGGGCTGATCAAGCCAACTATATGTTAGTTA 303
QY 1005 CAACCATCTG 1014
Db 304 CAACCATCTG 313

RESULT 9
US-10-424-599-70162

; Sequence 70162, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 70162
; LENGTH: 844
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34368C.1
US-10-424-599-70162

Query Match 12.1%; Score 185; DB 18; Length 844;
Best Local Similarity 100.0%; Pred. No. 2e-88;

Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1167 GCCTTACGTAGGGGATATGACAGAGGCCATGATGATGATCACTTCAGAGATATTCATGGG 1226
Db 9 GCCTTACGTAGGGGATATGACAGAGGCCATGATGATGATCACTTCAGAGATATTCATGGG 68
QY 1227 TGGAAAGAACACATTTGTTTGACACACATGTGAGGATTCCTTTAGCTGCTCTAT 1286
Db 69 TGGAAAGAACACATTTGTTTGACACACATGTGAGGATTCCTTTAGCTGCTCTAT 128
QY 1287 TATCTTGACCTTGTGCTCTTGTGAGCTGAGCACTGAAATCCAGTTTAAAGCTGAA 1346
Db 129 TATCTTGACCTTGTGCTCTTGTGAGCTGAGCACTGAAATCCAGTTTAAAGCTGAA 188
QY 1347 TGAGG 1351
Db 189 TGAGG 193

RESULT 10
US-10-424-599-116597

; Sequence 116597, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 116597
; LENGTH: 661
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(661)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_76298C.1
US-10-424-599-116597

Query Match 9.5%; Score 146; DB 18; Length 661;
Best Local Similarity 100.0%; Pred. No. 2e-67;

Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 ATGAAAATGTTCTCTTATTAATGAGACCTCAGAACACTTTTGTACAGAGGCTGATT 876
Db 56 ATGAAAATGTTCTCTTATTAATGAGACCTCAGAACACTTTTGTACAGAGGCTGATT 115
QY 877 GATCTTGCATCGCAGAGAACACTTTGATGAGTGAATGATCAAGATGCTCAGACC 936
Db 116 GATCTTGCATCGCAGAGAACACTTTGATGAGTGAATGATCAAGATGCTCAGACC 175
QY 937 AAAATGAATCTGTGTGCTGATT 962
Db 176 AAAATGAATCTGTGTGCTGATT 201

```
RESULT 11
US-10-025-003-5
; Sequence 5, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-5

Query Match
Best Local Similarity 100.0%; Score 146; DB 14; Length 1533;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 ATGGAATAATGTTCTTTCATTATATGGAAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
DB 817 ATGGAATAATGTTCTTTCATTATATGGAAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
QY 877 GATCTTGCATGCGGAGGAAACCTTGATTGTGTGAGATGACTTCAAGAGTGTCAGACC 936
DB 877 GATCTTGCATGCGGAGGAAACCTTGATTGTGTGAGATGACTTCAAGAGTGTCAGACC 936
QY 937 AAAATGAATCTGTGTGTTGATT 962
DB 937 AAAATGAATCTGTGTGTTGATT 962

RESULT 12
US-10-025-003-9
; Sequence 9, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-9

Query Match
Best Local Similarity 100.0%; Score 146; DB 14; Length 1533;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 ATGGAATAATGTTCTTTCATTATATGGAAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
DB 817 ATGGAATAATGTTCTTTCATTATATGGAAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
QY 877 GATCTTGCATGCGGAGGAAACCTTGATTGTGTGAGATGACTTCAAGAGTGTCAGACC 936
DB 877 GATCTTGCATGCGGAGGAAACCTTGATTGTGTGAGATGACTTCAAGAGTGTCAGACC 936
QY 937 AAAATGAATCTGTGTGTTGATT 962
DB 937 AAAATGAATCTGTGTGTTGATT 962

RESULT 13
US-10-718-952-5
; Sequence 5, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-5

Query Match
Best Local Similarity 100.0%; Score 146; DB 19; Length 1533;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 ATGGAATAATGTTCTTTCATTATATGGAAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
DB 817 ATGGAATAATGTTCTTTCATTATATGGAAGCCCTCAGAACACTTTTGTACAGGCGTGATT 876
QY 877 GATCTTGCATGCGGAGGAAACCTTGATTGTGTGAGATGACTTCAAGAGTGTCAGACC 936
DB 877 GATCTTGCATGCGGAGGAAACCTTGATTGTGTGAGATGACTTCAAGAGTGTCAGACC 936
QY 937 AAAATGAATCTGTGTGTTGATT 962
DB 937 AAAATGAATCTGTGTGTTGATT 962

RESULT 14
US-10-718-952-9
; Sequence 9, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
```

; PRIOR FILING DATE: APRIL 8, 1997
 ; PRIOR APPLICATION NUMBER: PCT/US98/068322
 ; PRIOR FILING DATE: APRIL 7, 1998
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 9
 ; LENGTH: 1533
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; US-10-718-952-9

Query Match	9.5%;	Score 146;	DB 19;	length 1533;
Best Local Similarity	100.0%;	Pred. No. 2.1e-67;		
Matches 146;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

RESULT 15
US-10-025

```

US-10-025-003-1
; Sequence 1, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025.003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR FILING DATE: APRIL 7, 1998
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-1

```

Query Match 9.5%; Score 146; DB 14; Length 1760;
Best Local Similarity 100.0%; Pred. No. 2.1e-67;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	81	ATGGAAGAAATGTCCTTTCAATTAATGGAACCCCTCAGAACACTTTTGTATCAGGGCTGATT	876
	870	ATGGAAGAAATGTCCTTTCAATTAATGGAACCCCTCAGAACACTTTTGTATCAGGGCTGATT	929
Db	877	GATCTTGCCATCGCGAGGAACACTTTGATTGTGTGAGATGACTTCAAGAGTGTACAGC	936
QY	930	GATCTTGCCATCGCGAGGAACACTTTGATTGTGTGAGATGACTTCAAGAGTGTACAGC	989
Db	937	AAATGGAATCTGTGTTGGTTGATT	962
QY	990	AAATGGAATCTGTGTTGGTTGATT	1015

Search completed: June 8, 2005, 19:34:53
Job time : 951.618 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:47:55 ; Search time 92 Seconds
(without alignments)
2143.998 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 510
Sequence: 1 MFIEFKESNVKYTEI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	100.0	510	ADQ14501	Adq14501 Mutant so
2	510	100.0	510	ADS82002	Ads82002 Soybean m
3	423	82.9	510	ADQ14503	Adq14503 Mutant so
4	423	82.9	510	ADQ14505	Adq14505 Wild type
5	423	82.9	510	ADS82004	Ads82004 Soybean m
6	423	82.9	510	ADS82006	Ads82006 Soybean m
7	236	46.3	510	AAW79740	AAw79740 Soybean w
8	236	46.3	510	ADQ14499	Adq14499 Mutant so
9	236	46.3	510	ADQ14491	Adq14491 Wild type
10	236	46.3	510	ADS82000	Ads82000 Soybean m
11	236	46.3	510	ADS81994	Ads81994 Soybean m
12	220	43.1	510	AAW79741	AAw79741 Soybean m
13	220	43.1	510	ADQ14495	Adq14495 Mutant so
14	220	43.1	510	ADS81998	Ads81998 Soybean m
15	149	14.9	510	AAW7882	AAw7882 Maize myo
16	149	14.9	510	AAW96259	AAw96259 Porterei
17	149	14.9	512	ADP43919	ADp43919 Porterei
18	149	14.9	536	AAW24477	AAw24477 Nicotiana
19	67	13.1	380	AAQ32503	AAq32503 Arabidops
20	67	13.1	446	AAQ32502	AAq32502 Arabidops
21	67	13.1	446	AAQ09862	AAq09862 Arabidops
22	67	13.1	505	ADG73738	ADg73738 Myo-inos
23	67	13.1	510	AAQ50575	AAq50575 Arabidops
24	67	13.1	510	AAQ48935	AAq48935 Brassica
25	67	13.1	510	AAQ47286	AAq47286 MIP synth

26	67	13.1	510	8	ADN73525	Adh73525 Thale cre
27	67	13.1	510	8	ADP43920	ADp43920 Rice L-my
28	67	13.1	511	3	AAQ09861	AAq09861 Arabidops
29	67	13.1	511	3	AAQ32501	AAq32501 Arabidops
30	67	13.1	534	3	AAQ09860	AAq09860 Arabidops
31	67	13.1	581	3	AAQ50574	AAq50574 Arabidops
32	67	13.1	645	3	AAQ50573	AAq50573 Arabidops
33	44	8.6	394	5	ABP51436	ABp51436 Human MDD
34	32	6.3	98	3	AAQ44825	AAq44825 Zea maye
35	29	5.7	565	4	ABB65497	ABb65497 Drosophi1
36	20	3.9	45	3	AAQ50576	AAq50576 Arabidops
37	20	3.9	78	5	ABP07716	ABp07716 Human ORF
38	20	3.9	113	3	AAQ12422	AAq12422 Zea maye
39	20	3.9	128	3	AAQ32705	AAq32705 Zea maye
40	20	3.9	130	3	AAQ32704	AAq32704 Zea maye
41	20	3.9	138	3	AAQ19004	AAq19004 Zea maye
42	20	3.9	153	3	AAQ32703	AAq32703 Zea maye
43	20	3.9	174	3	AAQ19003	AAq19003 Zea maye
44	19	3.7	251	3	AAQ52875	AAq52875 Arabidops
45	19	3.7	316	3	AAQ52874	AAq52874 Arabidops

ALIGNMENTS

RESULT 1	
ADQ14501	ADQ14501 standard; protein; 510 AA.
XX	
AC	ADQ14501;
XX	
DT	23-SEP-2004 (first entry)
XX	
DE	Mutant soybean myo-inositol 1-phosphate synthase polypeptide #3.
XX	
KW	soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KW	raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
KW	mutant; mutein.
XX	
OS	Glycine max.
OS	Synthetic.
XX	
PN	US2004128713-A1.
XX	
PD	01-JUL-2004.
XX	
PF	21-NOV-2003; 2003US-00718952.
XX	
PR	08-APR-1997; 97US-00835751.
PR	07-APR-1998; 98WO-US0006822.
PR	26-APR-1999; 99US-0029315.
XX	
XX	11-MAR-2002; 2002US-00025003.
PA	(HITZ/) HITZ W D.
PA	(SEBA/) SEBASTIAN S A.
PA	(GRAC/) GRACE D J.
PA	(STRE/) STREIT L G.
PI	Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX	
DR	WPI: 2004-53135/51.
XX	
DR	N-PSDB; ADQ14500.
XX	
PT	New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT	useful for producing plants with decreased raffinose, stachyose, and
PT	phytic acid and increased sucrose, leading to valuable and useful soybean
PT	products.
XX	
PS	Claim 9, SEQ ID NO 12; 48pp; English.
XX	
CC	The invention relates to a nucleic acid fragment encoding a soybean myo-
CC	inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC	synthase having a decreased capacity for the synthesis of myo-inositol 1-

CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.

SQ Sequence 510 AA;

Query Match 100.0%; Score 510; DB 8; Length 510;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MFIENFKYSPNVKKTETETIISVNYETTELVHENRNGTYQMIYKPSVNYQFKTNTHVP 60
DB 1 MFIENFKYSPNVKKTETETIISVNYETTELVHENRNGTYQMIYKPSVNYQFKTNTHVP 60
QY 61 KLGVWLVGMGNGNSTLTGVTIANREDISWATKDKIQQANFGSLTQASAIRVGSFOGEE 120
DB 61 KLGVWLVGMGNGNSTLTGVTIANREDISWATKDKIQQANFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLPLPVNPPDIVFGGWDISNMNLADAMARAKVFDIDLOQLRPPYESSWPLPGI 180
DB 121 IYAPFKSLPLPVNPPDIVFGGWDISNMNLADAMARAKVFDIDLOQLRPPYESSWPLPGI 180
QY 131 YDPDPIAANOERRANNVTKTQEOBOYQIIKDIKAFKERTKDKVYVMTANTERYSNLV 240
DB 131 YDPDPIAANOERRANNVTKTQEOBOYQIIKDIKAFKERTKDKVYVMTANTERYSNLV 240
QY 241 VGLNPTMENTLAAVNRNEAISPSTLYAIACMEVNPFIKNSPONTFPGGLDILAIANT 300
DB 241 VGLNPTMENTLAAVNRNEAISPSTLYAIACMEVNPFIKNSPONTFPGGLDILAIANT 300
QY 301 LIIGDDFKSGQTKMSVLVDFLVGAGIKRPTISVYNHLGNDGNMISAPOFRSKEISKS 360
DB 301 LIIGDDFKSGQTKMSVLVDFLVGAGIKRPTISVYNHLGNDGNMISAPOFRSKEISKS 360
QY 361 NVVDMMVNSNALIYEPGEHPDHVVVTKYVPYVGDSCRAMDEYTSBIFMGKNTIVLHNTC 420
DB 361 NVVDMMVNSNALIYEPGEHPDHVVVTKYVPYVGDSCRAMDEYTSBIFMGKNTIVLHNTC 420
QY 421 EDSLLAAPIIDLVLLAELSTRIOFKAENEGKFSFHPVATILSYLTAKAPLVPGTPPVN 480
DB 421 EDSLLAAPIIDLVLLAELSTRIOFKAENEGKFSFHPVATILSYLTAKAPLVPGTPPVN 480
QY 481 ALSKORAMLENIMRAACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRAACVGLAPENNMILEYK 510
```

RESULT 2

ADS82002 standard; protein; 510 AA.

ADS82002;

18-NOV-2004 (first entry)

Soybean myo-inositol 1-phosphate synthase mutant #2.

CC Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; mutein;
CC phytic acid; raffinose; stachyose; sucrose; inorganic phosphate;
CC flaculence; mutant.

OS Glycine max; line 29010CP01.
OS Synthetic.

XX Key Location/Qualifiers
FH Misc-difference 87 /note= "Wild-type Gly substituted by Asp"

US2003074685-A1.

17-APR-2003.

11-MAR-2002; 2002US-00025003.

08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US0006822.

XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.

Hitz WD, Sebastian SA;

WPI, 2004-639957/62.

N-PSDB; ADS82001.

DR Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX phytic acid and inorganic phosphate content of soybean seeds.

Claim 9; SEQ ID NO 12; 34bp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with a
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flaculence
CC when consumed by humans. The present sequence represents a mutant myo-
CC inositol 1-phosphate synthase.

SQ Sequence 510 AA;

Query Match 100.0%; Score 510; DB 8; Length 510;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MFIENFKYSPNVKKTETETIISVNYETTELVHENRNGTYQMIYKPSVNYQFKTNTHVP 60
DB 1 MFIENFKYSPNVKKTETETIISVNYETTELVHENRNGTYQMIYKPSVNYQFKTNTHVP 60
QY 61 KLGVWLVGMGNGNSTLTGVTIANREDISWATKDKIQQANFGSLTQASAIRVGSFOGEE 120
```

Dd	61	KLGVNLVHGNGNGSTLTGCVLANRBDISWATKDXIQOANYFESLTOASAIRVGSFOGEE	120
Qy	121	IYAPFKSLPMPNPDIVFGWDISNMNLADAMARAKVEDIDQOKLRPMESMVPJGI	180
Dd	121	IYAPFKSLPMPNPDIVFGWDISNMNLADAMARAKVEDIDQOKLRPMESMVPJGI	180
Qy	181	YDPDEIAANOEBRANNVIKGTQOEYOQOIIKDIKAFKATKDVKVVMLTANERYSNLV	240
Dd	181	YDPDEIAANOEBRANNVIKGTQOEYOQOIIKDIKAFKATKDVKVVMLTANERYSNLV	240
Qy	241	VGANDTMENLAAVDRENEAISEPSTLYAIACWENVPFINGSPONTFVPGJLDLAIARNT	300
Dd	241	VGANDTMENLAAVDRENEAISEPSTLYAIACWENVPFINGSPONTFVPGJLDLAIARNT	300
Qy	301	LIGDDPFKSGQTKMKSVLVDPLVAGIKPTPSIVSYNHLGNNDGMNLSAPQTPRSKEISKS	360
Dd	301	LIGDDPFKSGQTKMKSVLVDPLVAGIKPTPSIVSYNHLGNNDGMNLSAPQTPRSKEISKS	360
Qy	361	NVVDOMVNSNALIYEPGEHPDHVVYIKYIPYVDSGRANDEYSEIFMGSKNTIYLHNTC	420
Dd	361	NVVDOMVNSNALIYEPGEHPDHVVYIKYIPYVDSGRANDEYSEIFMGSKNTIYLHNTC	420
Qy	421	EDSLLAAPILDLVLLAEISTRIOFQAEHBGKHSHPVATIIISYLTKAPLVPFGTPVYN	480
Dd	421	EDSLLAAPILDLVLLAEISTRIOFQAEHBGKHSHPVATIIISYLTKAPLVPFGTPVYN	480
Qy	481	ALSKORAMENIMRACVGLAPENNMILEXX	510
Dd	481	ALSKORAMENIMRACVGLAPENNMILEXX	510

RESULT 3
ADQ14503
ID ADQ14503 standard; protein; 510 AA.

AC ADQ14503;

DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #4.

KM Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KM raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
KM mutant; mutein.

OS	Glycine max.
OS	Synthetic.

PN US2004128713-A1.

01-JUL-2004.
PD

PF 21-NOV-2003; 2003US-00718952.

PR 08-APR-1997; 97US-00835751.

PR 26-APR-1999; 99US-00299315.

XX

PA (SEBA/) SEBASTIAN S A.

PA (STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

DR WPI; 2004-533135/51.

XX

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phylic acid and increased sucrose, leading to valuable and useful soybean
PT products.

XX Example 8; SEQ ID NO 14; 48pp; English.

XX

XX The invention relates to a nucleic acid fragment encoding a soybean myo-

CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate

CC synthase having a decreased capacity for the synthesis of myo-inositol 1-

CC phosphate. The invention also relates to a chimeric gene operably linked

CC to suitable regulatory sequences, where expression of the chimeric gene

CC results in a decrease in expression of an endogenous or native gene

CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant

CC comprising the chimeric gene, a method of making the soybean plant, a

CC seed of the soybean plant, a soy protein product derived from the

CC processing of soybean seeds, a method of making or producing a soy

CC protein product and a method of using a soybean plant homozygous for at

CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having

CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The

CC nucleic acid fragment and methods are useful for producing plants with

CC decreased raffinose, stachyose and phytyl acid content and increased

CC sucrose and inorganic phosphate content, leading to valuable and useful

CC soybean products. This sequence represents a mutant soybean myo-inositol

CC 1-phosphate synthase polypeptide of the invention.

Sequence 510 AA;

Query Match 82.9%; Score 423; DB 8; Length 510;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ISWATKDKIQQANYFGSLTQASAIRVGSFQGEIYAPFKSLPMVNPDDIVFGWDISM 147

Db 88 ISWATKDKIQQANYFGSLTQASAIRVGSFGGEIYAPFKSLPWNPPDIVFGGWDISM 147

QY 148 NLADAMARAKVFDIDLQKOLRPMESMVLPGIYDPDFIAANQERANNVIKGTKQEQVQ 207

Db 148 NLADAMARAKVEDIDLQKQLRPYMESWPLPGIYDPDFIAAQEERANNVIKGTKEQVQ 207

QY 208 QIKDIKAFKATKVDKVVLTANTERYSNLVGLNDTMENLLAAVDRNEAEISPSTLY 267

Db 208 QIKDIKAFKATKVDKVVLTANTERYSNLVGGLNDTMENLLAADVDRNEAISPSTLY 267

QY 268 AIACMENVPFINGSPONTFVPGILDLAIARNTLIGDDFKSGQTKMSVLVDFLVGAGI 327

Db 268 AIACMENVPFINGSPQNTFVPGILDIAIRNTLIGDDFKSGQTKMKSVLVDFVGAGI 327

QY 328 KPTSVSYNHLGNNDGMNLSAPQTFRSKEISKSNVDDMVNSNAILYEPGEHPDHVVIK 387

Db 328 KPTSI VSYNHLGNNDGM LSA P Q TFRSKEISKSNVDDMVNSNAILYEPGEHPDVVVK 387

QY 388 YVPYVGDSKRAMDEYTS E I FMGKNTIVLHNTCEDSLAP I I L D V L L A E L S T R I Q F K A 447

Db 388 YVPYVGDSKRAMDEYTS EIFMGKNTIVLHNTCEDSLA P I I D L V L A E L S T R I Q F K A 447

QY 448 ENECKFHSFHPVATILSYLTKAPLVPPGTPVNALSKQAMLENIMRACVGLAPENMIL 507

Db 448 ENEGKFHSFHVATILSYLTAPLVPPGTPVVALSKQAMLENIMRACVGLAPENNML 507

QY 508 EYK 510

Db 508 EYK 510

ADQ14505

XX

XX

XX

11

KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.

```

XX Glycine max.
OS US2004128713-A1.
PN 01-JUL-2004.
PF 21-NOV-2003; 2003US-00718952.
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-00293315.
PR 11-MAR-2002; 2002US-00025003.
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX MPI: 2004-533135/51.
XX N-PSDB; ADQ14504.
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX useful for producing plants with decreased raffinose, stachyose, and
XX phytic acid and increased sucrose, leading to valuable and useful soybean
XX products.
XX Claim 3; SEQ ID NO 16; 48pp; English.
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents a wild type soybean myo-
XX inositol 1-phosphate synthase polypeptide of the invention.
SQ Sequence 510 AA;
Query Match 82.9%; Score 423; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 328 KPTSIIVSYNHLGNDGNMLSAPOTFRSKEISKSNVVDWMVNSNALYEPGEHPDHVVIX 387
QY 388 YVPYVGDSEKRAMDEXTSEIFMGKNTIYLAHNTCEDSLAAPIIIDLVLAELSTRIOFKA 447
DB 388 YVPYVGDSEKRAMDEXTSEIFMGKNTIYLAHNTCEDSLAAPIIIDLVLAELSTRIOFKA 447
QY 448 ENEGKFHSHFHVATILSYLTKAPLVPPEPTPVVNALSKORAMLENIMRACVGIABENNMI 507
DB 448 ENEGKFHSHFHVATILSYLTKAPLVPPEPTPVVNALSKORAMLENIMRACVGIABENNMI 507
QY 508 EYK 510
DB 508 EYK 510
RESULT 5
ADSS82004
ID ADSS82004 standard; protein; 510 AA.
AC ADSS82004;
DT 18-NOV-2004 (first entry)
DE Soybean myo-inositol 1-phosphate synthase wild-type 4.
XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX Glycine max; line 290180P03.
XX US2003074685-A1.
XX 17-APR-2003.
XX 11-MAR-2002; 2002US-00025003.
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX Hitz WD, Sebastian SA;
XX MPI: 2004-639957/62.
XX N-PSDB; ADS82003.
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX phytic acid and inorganic phosphate content of soybean seeds.
XX Example 8; SEQ ID NO 14; 34pp; English.
XX The invention relates to an isolated nucleic acid fragment encoding a
XX soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX phosphate synthase having decreased capacity for the synthesis for myo-
XX inositol 1-phosphate. Also included are a chimeric gene (comprising the
XX nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX complement, subfragment or the complement of the subfragment, operably
XX linked to suitable regulatory sequences, where expression of the chimeric
XX gene results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX comprising the chimeric gene (with a heritable phenotype of a seed phytic
XX acid content of less than 17 micromol/g, a seed content of raffinose plus
XX stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX greater than 200 micromol/g), provided that the plant is not LR33), seeds
XX from the plant, making a soybean plant with the heritable phenotype
XX (comprising crossing LR33 or the plant comprising the chimeric gene with
XX an elite soybean plant and selecting a progeny plant of the cross of
XX crossing step that has a heritable phenotype as mentioned above), seeds
XX of soybean plant made by the above method, a soy protein product derived
XX from seeds of a soybean plant (homozygous for one or more gene encoding a

```

CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with
 CC a heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence represents a wild-type myo-
 CC inositol 1-phosphate synthase.

Sequence 510 AA:

Query Match 82.9%; Score 423; DB 8; Length 510;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ISMATKDKIQOANVFGSLTQASAIRVSGFGEETIAPRKSLLPMVNPDDIVFGWDISNM 147
 DB 88 ISMATKDKIQOANVFGSLTQASAIRVSGFGEETIAPRKSLLPMVNPDDIVFGWDISNM 147
 QY 148 NLADAMARAKVFDIDLQKQLRPYMESWVPLPGIYDPDFIAANOBERANNAVIGTKQOEVOQ 207
 DB 148 NLADAMARAKVFDIDLQKQLRPYMESWVPLPGIYDPDFIAANOBERANNAVIGTKQOEVOQ 207
 QY 208 QIIKDIKAFKEATKYDKVVLMTANTERYSNLVGINDTMENLLAADRNEAISPSTLY 267
 DB 208 QIIKDIKAFKEATKYDKVVLMTANTERYSNLVGINDTMENLLAADRNEAISPSTLY 267
 QY 268 AIACMVENVFPIINGSPQNTFVPGGLDLAIARNTLLIGDDPFKSGQTKMSVLDVFLVAGAI 327
 DB 268 AIACMVENVFPIINGSPQNTFVPGGLDLAIARNTLLIGDDPFKSGQTKMSVLDVFLVAGAI 327
 QY 328 KPTSIIVSYNHLGNNDGNMLSAPQTFRSKESISKNVVDVDMVSNAILYEPGHPDHVVVIX 387
 DB 328 KPTSIIVSYNHLGNNDGNMLSAPQTFRSKESISKNVVDVDMVSNAILYEPGHPDHVVVIX 387
 QY 388 YVPYVGDSCRANDEYTSIEIFMGCKNTIVLANTCEBSLLAAPILLDLVLAELSTRIQFKA 447
 DB 388 YVPYVGDSCRANDEYTSIEIFMGCKNTIVLANTCEBSLLAAPILLDLVLAELSTRIQFKA 447
 QY 448 ENEGKFSHPHVAITLSTYTKAPLVPCGPVYNALSKORAMLENTMRACVGLAPENMML 507
 DB 448 ENEGKFSHPHVAITLSTYTKAPLVPCGPVYNALSKORAMLENTMRACVGLAPENMML 507
 QY 508 EYK 510
 DB 508 EYK 510

RESULT 6
 ADS82006
 ID ADS82006 standard; protein; 510 AA.
 AC ADS82006;
 DT 18-NOV-2004 (first entry)
 XX
 XX
 DE Soybean myo-inositol 1-phosphate synthase wild-type 2.
 XX
 XX
 KM Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
 KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
 XX
 OS Glycine max; cultivar Wye.
 XX
 PN US2003074685-A1.
 PD 17-APR-2003.
 PF 11-MAR-2002; 2002US-00025003.
 PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.
 XX
 XX (HITZ/) HITZ M D.
 PA (SEBA/) SEBASTIAN S A.
 XX
 XX
 PI Hitz WD, Sebastian SA;
 XX
 XX WPI. 2004-639957/62.
 DR N-PSDB; ADS82005.
 XX
 XX
 PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
 PT phytic acid and inorganic phosphate content of soybean seeds.
 PS Claim 3, SEQ ID NO 16; 34pp; English.

The invention relates to an isolated nucleic acid fragment encoding a
 CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 CC phosphate synthase having decreased capacity for the synthesis for myo-
 CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
 CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 CC complement, subfragment or the complement of the subfragment, operably
 CC linked to suitable regulatory sequences, where expression of the chimeric
 CC gene results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
 CC acid content of less than 17 micromol/g, a seed content of raffinose plus
 CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
 CC greater than 200 micromol/g), provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype
 CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homozygous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with
 CC a heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence represents a wild-type myo-
 CC inositol 1-phosphate synthase.

Sequence 510 AA:

Query Match 82.9%; Score 423; DB 8; Length 510;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ISMATKDKIQOANVFGSLTQASAIRVSGFGEETIAPRKSLLPMVNPDDIVFGWDISNM 147
 DB 88 ISMATKDKIQOANVFGSLTQASAIRVSGFGEETIAPRKSLLPMVNPDDIVFGWDISNM 147
 QY 148 NLADAMARAKVFDIDLQKQLRPYMESWVPLPGIYDPDFIAANOBERANNAVIGTKQOEVOQ 207
 DB 148 NLADAMARAKVFDIDLQKQLRPYMESWVPLPGIYDPDFIAANOBERANNAVIGTKQOEVOQ 207
 QY 208 QIIKDIKAFKEATKYDKVVLMTANTERYSNLVGINDTMENLLAADRNEAISPSTLY 267
 DB 208 QIIKDIKAFKEATKYDKVVLMTANTERYSNLVGINDTMENLLAADRNEAISPSTLY 267
 QY 268 AIACMVENVFPIINGSPQNTFVPGGLDLAIARNTLLIGDDPFKSGQTKMSVLDVFLVAGAI 327
 DB 268 AIACMVENVFPIINGSPQNTFVPGGLDLAIARNTLLIGDDPFKSGQTKMSVLDVFLVAGAI 327
 QY 328 KPTSIIVSYNHLGNNDGNMLSAPQTFRSKESISKNVVDVDMVSNAILYEPGHPDHVVVIX 387
 DB 328 KPTSIIVSYNHLGNNDGNMLSAPQTFRSKESISKNVVDVDMVSNAILYEPGHPDHVVVIX 387

QY 388 YVPYVGDSSGRANDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILLDLVLAELSTRIOFKA 447
DB 388 YVPYVGDSSGRANDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILLDLVLAELSTRIOFKA 447
QY 448 ENEGKFSHFHPVATITISYLTAKA PLVPRGPRVYNALSKORAMLENTMRACVGLAPNNMIL 507
DB 448 ENEGKFSHFHPVATITISYLTAKA PLVPRGPRVYNALSKORAMLENTMRACVGLAPNNMIL 507
QY 508 EYK 510
DB 508 EYK 510

RESULT 7
AAW79740
ID AAW79740 standard; protein; 510 AA.

AC AAW79740;
DT 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)

DE Soybean wild-type myo-inositol 1-phosphate synthase.

KW Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KW phytic acid.

OS Glycine max; line LR13.

PN W09845448-A1.

PD 15-OCT-1998.

PF 07-APR-1998; 98WO-US006822.

PR 08-APR-1997; 97US-00835751.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Hitz WD, Sebastian SA;

PR WPI: 1998-568353/48.

DR N-PSDB; AAV62440.

PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
PT phytic acid, etc.

PS Example 5; Page 45-47; 63pp; English.

CC This is the amino acid sequence of soybean myo-inositol 1-phosphate
CC synthase (MI 1-PS) deduced from the coding region of an isolated cDNA
CC clone (see AAV62440). MI 1-PS is involved in glucose metabolism to phytic
CC acid, raffinose and stachyose. A mutant MI 1-PS (see AAW79741) has been
CC identified in soybean line LR13, a mutagenised line of low raffinose
CC saccharide phenotype. Sequencing revealed a single base change mutation
CC in the LR13 gene sequence that resulted in a K396N substitution in the
CC mutant protein. The mutation results in a seed phenotype of very low
CC raffinose saccharide sugars, very high sucrose and low phytic acid. The
CC mutated nucleic acid is used to alter the raffinose saccharide, sucrose,
CC phytic acid and inorganic phosphate content of soybean seeds, leading to
CC useful soybean products, e.g. a seed phytic acid content of less than 17
CC ug/g, a seed content of raffinose and stachyose combined of less than
CC 14.5 ug/g, and a seed sucrose content greater than 200 ug/g. (Updated on
CC 17-OCT-2003 to standardise OS field)

XX Sequence 510 AA;

Query Match 46.3%; Score 236; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.5e-227;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPDFIAANOEBRANNVIKGTKEOVQOIIKIDIKAFKATKVDKVVVLTWANTER 235

DB 176 PLPGIYDPDFIAANOEBRANNVIKGTKEOVQOIIKIDIKAFKATKVDKVVVLTWANTER 235
QY 236 YSNLVAGNDTMENILAAVDNBEAIESPTLYALACWENVPFINGSPONTFVEGLIDLA 295
DB 236 YSNLVAGNDTMENILAAVDNBEAIESPTLYALACWENVPFINGSPONTFVEGLIDLA 295
QY 296 IARNTLIGDDFKSGQTKMSVLVDFLVGAGIKPSTISVYNHLGNDGMNISAPQTERSK 355
DB 296 IARNTLIGDDFKSGQTKMSVLVDFLVGAGIKPSTISVYNHLGNDGMNISAPQTERSK 355
QY 356 EISKSNVVDMDVNSNALIYEPGEHPDHVVYIKYVPYVGDSSGRANDEYTSSEIFMGSK 411
DB 356 EISKSNVVDMDVNSNALIYEPGEHPDHVVYIKYVPYVGDSSGRANDEYTSSEIFMGSK 411

RESULT 8
ADQ14499
ID ADQ14499 standard; protein; 510 AA.

AC ADQ14499;

DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #2.

XX Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;

KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;

KW mutant; mutein.

OS Glycine max.

OS Synthetic.

PN US2004128713-A1.

PD 01-JUL-2004.

PF 21-NOV-2003; 2003US-00718952.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.

PR 26-APR-1999; 99US-00299315.

PR 11-MAR-2002; 2002US-00025003.

PA (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.

PA (STRE/) STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

PR WPI: 2004-533135/51.

DR N-PSDB; ADQ14498.

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

PS Example 8; SEQ ID NO 10; 48pp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having

CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.

XX
SQ Sequence 510 AA;

Query Match 46.3%; Score 236; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.5e-227;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPPFLAANOEERANNVYKGTQKQOVQIIKDIAKFEATKVDKVVVLTMTANTER 235
DB 176 PLPGIYDPPFLAANOEERANNVYKGTQKQOVQIIKDIAKFEATKVDKVVVLTMTANTER 235
QY 236 YSNLVVGLNDTMENTLLAAVDRNEAEISPSSTLYAIAACMENVPFINGSPONTFVPGIIDA 295
DB 236 YSNLVVGLNDTMENTLLAAVDRNEAEISPSSTLYAIAACMENVPFINGSPONTFVPGIIDA 295
QY 296 IARNTLIGDDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSK 355
DB 296 IARNTLIGDDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSK 355
QY 356 EISKSNVVDVMNSNAILYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSEIFMGCK 411
DB 356 EISKSNVVDVMNSNAILYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSEIFMGCK 411

RESULT 9
ADQ14491
ID ADQ14491 standard; protein; 510 AA.

XX ADQ14491;
DT 23-SEP-2004 (first entry)

XX Wild type soybean myo-inositol 1-phosphate synthase polypeptide #1.

KW Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KM raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.

XX Glycine max.

XX US2004128713-A1.

XX 01-JUL-2004.

XX 21-NOV-2003; 2003US-00718952.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX 06-APR-1999; 99US-00299315.

XX 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX (GRAC/) GRACE D J.

XX (STRE/) STREIT L G.

XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;

XX WPI; 2004-533135/51.

XX N-PSDB; ADQ14490.

XX Claim 3; SEQ ID NO 2; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a wild type soybean myo-
CC inositol 1-phosphate synthase polypeptide of the invention.

XX
SQ Sequence 510 AA;

Query Match 46.3%; Score 236; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.5e-227;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPPFLAANOEERANNVYKGTQKQOVQIIKDIAKFEATKVDKVVVLTMTANTER 235
DB 176 PLPGIYDPPFLAANOEERANNVYKGTQKQOVQIIKDIAKFEATKVDKVVVLTMTANTER 235
QY 236 YSNLVVGLNDTMENTLLAAVDRNEAEISPSSTLYAIAACMENVPFINGSPONTFVPGIIDA 295
DB 236 YSNLVVGLNDTMENTLLAAVDRNEAEISPSSTLYAIAACMENVPFINGSPONTFVPGIIDA 295
QY 296 IARNTLIGDDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSK 355
DB 296 IARNTLIGDDDFKSGQTKMKSVLVDFLVGAGIKPTISIVSYNHLGNNDGNLSAPQTFRSK 355
QY 356 EISKSNVVDVMNSNAILYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSEIFMGCK 411
DB 356 EISKSNVVDVMNSNAILYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSEIFMGCK 411

RESULT 10
ADS82000
ID ADS82000 standard; protein; 510 AA.

XX ADS82000;

XX 18-NOV-2004 (first entry)

XX Soybean myo-inositol 1-phosphate synthase wild-type 3.

XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;

KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.

XX Glycine max; line 29004UP01.

XX US2003074685-A1.

XX 17-APR-2003.

XX 11-MAR-2002; 2002US-00025003.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX Hitz WD, Sebastian SA;

XX WPI; 2004-639957/62.

XX N-PSDB; ADS81999.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

PS Example 8, SEQ ID NO 10; 34pp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a wild-type myo-inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 46.3%; Score 236; DB 8; Length 510;

Best Local Similarity 100.0%; Pred. No. 4.5e-227; Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDDPFIANAQBERANNVYKGTQKQOVQOIIKDIKAFKATKVDKVVVLTANTER 235
DB 176 PLPGIYDDPFIANAQBERANNVYKGTQKQOVQOIIKDIKAFKATKVDKVVVLTANTER 235
QY 236 YSNLVVGLNDTMENTLLAAVDNRNEAISPSTLYAIAACWENVPFINGSPQNTFVGLIDLA 295
DB 236 YSNLVVGLNDTMENTLLAAVDNRNEAISPSTLYAIAACWENVPFINGSPQNTFVGLIDLA 295
QY 236 IARNTLIGDDDFKSGQTKRKSVLVDFLVGAGIKPTSIYSYNHLGNDGNMNSAPQTFPSK 355
DB 236 IARNTLIGDDDFKSGQTKRKSVLVDFLVGAGIKPTSIYSYNHLGNDGNMNSAPQTFPSK 355
QY 356 EISKSNVDDMNSAAILYEPGEHPDHVVVYKVPYVGSKRAMEYTSIFMGSK 411
DB 356 EISKSNVDDMNSAAILYEPGEHPDHVVVYKVPYVGSKRAMEYTSIFMGSK 411

RESULT 11

ADS81994 ID ADS81994 standard; protein; 510 AA.

XX ADS81994;

DT 18-NOV-2004 (first entry)

DE Soybean myo-inositol 1-phosphate synthase wild-type 1.

XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.

XX Glycine max; cultivar Wye.

XX US2003074685-A1.

XX 17-APR-2003.

XX 11-MAR-2002; 2002US-00025003.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX (HITZ/) HITZ M D.
XX (SEBA/) SEBASTIAN S A.

XX Hitz MD, Sebastian SA;

XX WPI, 2004-639957/62.

XX N-PsDB; ADS81993.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

XX Claim 3, SEQ ID NO 2; 34pp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a wild-type myo-inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 46.3%; Score 236; DB 8; Length 510;

Best Local Similarity 100.0%; Pred. No. 4.5e-227; Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDDPFIANAQBERANNVYKGTQKQOVQOIIKDIKAFKATKVDKVVVLTANTER 235
DB 176 PLPGIYDDPFIANAQBERANNVYKGTQKQOVQOIIKDIKAFKATKVDKVVVLTANTER 235

QY 236 YSNLVVGLNDTMENTLLAAVDNRNEAISPSTLYAIAACWENVPFINGSPQNTFVGLIDLA 295
DB 236 YSNLVVGLNDTMENTLLAAVDNRNEAISPSTLYAIAACWENVPFINGSPQNTFVGLIDLA 295

QY 236 IARNTLIGDDDFKSGQTKRKSVLVDFLVGAGIKPTSIYSYNHLGNDGNMNSAPQTFPSK 355
DB 236 IARNTLIGDDDFKSGQTKRKSVLVDFLVGAGIKPTSIYSYNHLGNDGNMNSAPQTFPSK 355


```

Db      296 IARNLTIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNLSAPQTFRSK 355
Qy      356 EISKSNVVDMDMNSNALIYEPGEHPDHVVVIKYVPVYVSGSKRAMDYTSEIFMGK 411
        |||
        |||
        |||
Db      356 EISKSNVVDMDMNSNALIYEPGEHPDHVVVIKYVPVYVSGSKRAMDEYTSIEFMGK 411

RESULT 12
AAM79741
ID      AAM79741 standard; protein; 510 AA.
AC      AAM79741;
XX
XX
DT      17-OCT-2003 (revised)
DT      02-FEB-1999 (first entry)
XX
XX
DE      Soybean mutant myo-inositol 1-phosphate synthase.
XX
XX      Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
XX      phytic acid.
XX
XX      Glycine max; line LR33.
XX
XX      WO9845448-A1.
XX
XX      15-OCT-1998.
XX
XX      07-APR-1998; 98WO-US006822.
XX
XX      08-APR-1997; 97US-00835751.
XX
XX      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX      Hitz WD, Sebastian SA;
XX      WPI; 1998-568353/48.
XX      N-PSDB; AAV62443.
XX
XX      Soybean plants containing altered myo-inositol-1-phosphate gene - useful
XX      for generating plants with altered levels of e.g. raffinose, stachyose,
XX      phytic acid, etc.
XX
XX      Example 5; Page 49-51; 63pp; English.
XX
XX      This is the amino acid sequence of a mutant soybean myo-inositol 1-
XX      phosphate synthase (MI 1-PS) deduced from the coding region of an
XX      isolated cDNA clone (see AAV62443). MI 1-PS is involved in glucose
XX      metabolism to phytic acid, raffinose and stachyose. The MI 1-PS was
XX      identified in soybean line LR33, a mutagenised line of low raffinose
XX      saccharide phenotype. Sequencing revealed a single base change mutation
XX      in the LR33 gene sequence that resulted in a K396N substitution in the
XX      mutant protein compared to wild-type MI 1-PS (see AAM79740). The mutation
XX      results in a seed phenotype of very low raffinose saccharide sugars, very
XX      high sucrose and low phytic acid. The mutated nucleic acid is used to
XX      alter the raffinose saccharide, sucrose, phytic acid and inorganic
XX      phosphate content of soybean seeds, leading to useful soybean products,
XX      e.g. a seed phytic acid content of less than 17 ug/g, a seed content of
XX      raffinose and stachyose combined of less than 14.5 ug/g, and a seed
XX      sucrose content greater than 200 ug/g. (Updated on 17-OCT-2003 to
XX      standardise OS field)
XX
XX      Sequence 510 AA;
XX
XX      Query Match      43.1%; Score 220; DB 2; Length 510;
XX      Best Local Similarity 100.0%; Pred. No. 4.7e-211;
XX      Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      176 PLPGIYDPRFIANQGERANNVIKGRKQOVQOIIDIAFKFATVDKVVVLTMTANTER 235
        |||
        |||
        |||
Db      176 PLPGIYDPRFIANQGERANNVIKGRKQOVQOIIDIAFKFATVDKVVVLTMTANTER 235
        |||
        |||
        |||
Qy      236 YSNLVGLNDTMENLLAAVDNRBAEISPTLYAIACMENVPFINGSPONTFVPGILDLA 295
        |||
        |||
        |||

```

```

Db      236 YSNLVGLNDTMENLLAAVDNRBAEISPTLYAIACMENVPFINGSPONTFVPGILDLA 295
Qy      296 IARNLTIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNLSAPQTFRSK 355
        |||
        |||
        |||
Db      296 IARNLTIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNLSAPQTFRSK 355

RESULT 13
ADQ14495
ID      ADQ14495 standard; protein; 510 AA.
XX
XX      ADQ14495;
XX
XX      23-SEP-2004 (first entry)
XX
XX      Mutant soybean myo-inositol 1-phosphate synthase polypeptide #1.
XX
XX      Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
XX      raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
XX      mutant; mutein.
XX
XX      Glycine max.
XX      Synthetic.
XX
XX      Key      Location/Qualifiers
XX      FH      Misc-difference 396
XX      FT      /note= "Wild type Lys substituted by Asn"
XX      PN      US2004128713-A1.
XX
XX      01-JUL-2004.
XX
XX      21-NOV-2003; 2003US-00718952.
XX
XX      08-APR-1997; 97US-00835751.
XX      PR      07-APR-1998; 98WO-US006822.
XX      PR      26-APR-1999; 99US-00299315.
XX      PR      11-MAR-2002; 2002US-00025003.
XX
XX      (HITZ/) HITZ W D.
XX      PA      (SEBA/) SEBASTIAN S A.
XX      PA      (GRAC/) GRACE D J.
XX      PA      (STRE/) STREIT L G.
XX
XX      Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX      WPI; 2004-533135/51.
XX      DR      N-PSDB; ADQ14494.
XX
XX      New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
XX      useful for producing plants with decreased raffinose, stachyose, and
XX      phytic acid and increased sucrose, leading to valuable and useful soybean
XX      products.
XX
XX      Claim 9; SEQ ID NO 6; 48bp; English.
XX
XX      The invention relates to a nucleic acid fragment encoding a soybean myo-
XX      inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX      synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX      phosphate. The invention also relates to a chimeric gene operably linked
XX      to suitable regulatory sequences, where expression of the chimeric gene
XX      results in a decrease in expression of an endogenous or native gene
XX      encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX      comprising the chimeric gene, a method of making the soybean plant, a
XX      seed of the soybean plant, a soy protein product derived from the
XX      processing of soybean seeds, a method of making or producing a soy
XX      protein product and a method of using a soybean plant homozygous for at
XX      least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX      decreased capacity for the synthesis of myo-inositol 1-phosphate. The

```

CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents a mutant soybean myo-inositol
CC 1-phosphate synthase polypeptide of the invention.

XX Sequence 510 AA;

Query Match 43.1%; Score 220; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.7e-211;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPPFIANAQEEERANNVIGKTQKQVQOQIIKDIKAFKATKVDKVVVLTANTER 235
DB 176 PLPGIYDPPFIANAQEEERANNVIGKTQKQVQOQIIKDIKAFKATKVDKVVVLTANTER 235
QY 236 YSNLVVGLNDTMENTLAAVDREAEISPSSTLYAACWENVPFINGSPONTFVPGILDA 295
DB 236 YSNLVVGLNDTMENTLAAVDREAEISPSSTLYAACWENVPFINGSPONTFVPGILDA 295
QY 296 IARNTLIGDDPFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNTLSAPOTFRSK 355
DB 296 IARNTLIGDDPFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNTLSAPOTFRSK 355
QY 356 EISKSNVVDVMVNSNAILYEPGEHPDHVVVTKVYPVGD 395
DB 356 EISKSNVVDVMVNSNAILYEPGEHPDHVVVTKVYPVGD 395

RESULT 14

ADSB1998
ID ADSB1998 standard; protein, 510 AA.

AC ADSB1998;
DT 18-NOV-2004 (first entry)

XX Soybean myo-inositol 1-phosphate synthase mutant #1.

XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; muten;
KW phytic acid; raffinose; stachyose; sucrose; inorganic phosphate;
KM flatulence; mutant.

XX Glycine max; line LR33.
OS Synthetic.

XX Key location/Qualifiers

FT Misc-difference 396 /note= "Wild-type Lys substituted by Arg"

PN US2003074685-A1.

PD 17-APR-2003.

PF 11-MAR-2002; 2002US-00025003.

PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.

PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.

PI Hitz WD, Sebastian SA;

DR WPI: 2004-639957/62.
DR N-PSDB; ADSB1997.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.

PS Claim 9, SEQ ID NO 6; 34pp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences), where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate), and making or producing a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence represents a mutant myo-
CC inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 43.1%; Score 220; DB 8; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.7e-211;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPPFIANAQEEERANNVIGKTQKQVQOQIIKDIKAFKATKVDKVVVLTANTER 235
DB 176 PLPGIYDPPFIANAQEEERANNVIGKTQKQVQOQIIKDIKAFKATKVDKVVVLTANTER 235
QY 236 YSNLVVGLNDTMENTLAAVDREAEISPSSTLYAACWENVPFINGSPONTFVPGILDA 295
DB 236 YSNLVVGLNDTMENTLAAVDREAEISPSSTLYAACWENVPFINGSPONTFVPGILDA 295
QY 296 IARNTLIGDDPFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNTLSAPOTFRSK 355
DB 296 IARNTLIGDDPFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNTLSAPOTFRSK 355
QY 356 EISKSNVVDVMVNSNAILYEPGEHPDHVVVTKVYPVGD 395
DB 356 EISKSNVVDVMVNSNAILYEPGEHPDHVVVTKVYPVGD 395

RESULT 15

AAW97882
ID AAW97882 standard; protein, 510 AA.

AC AAW97882;

DT 07-JUN-1999 (first entry)

XX Maize myo-inositol 1-phosphate synthase.

XX Myo-inositol 1-phosphate synthase; maize; corn; phytate; phytic acid;
KW transgenic plant; animal nutrition; feedstuff; food.

XX Zea mays.

XX WO9905298-A1.

PD 04-FEB-1999.

PF 17-JUL-1996; 98WO-US014657.
XX
PR 22-JUL-1997; 97US-0053371P.
PR 28-JUL-1997; 97US-0053944P.
PR 08-AUG-1997; 97US-0055526P.
PR 11-AUG-1997; 97US-0055446P.
PR 18-MAY-1996; 98US-0085852P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Martino-Catt SJ, Wang H, Beach LR, Bowen BA, Wang X;
XX
DR WPI: 1999-142948/12.
DR N-PSDB; AAX24407.
XX
XX New polynucleotides controlling phytate metabolism in plants - useful for
PT improving the nutritional content of plants, by enhancing levels of non-
PT phytate phosphorus, and reducing phytate levels.
XX
PS Claim 1a; Page 77-78; 86pp; English.
XX
CC This is the amino acid sequence of maize myo-inositol 1-phosphate
CC synthase, an enzyme involved in the metabolism of phytate. cDNA (see
CC AAX24407) encoding the enzyme was isolated from a maize embryo (15 day
CC post-pollination) cDNA library. Polynucleotides (see AAX24400, AAX24403,
CC AAX24407 and AAX24410-12) encoding maize phosphatidylinositol-3-kinase
CC (see AAW97880), myo-inositol 1,3,4-trisphosphate 5/6-kinase (see
CC AAW97881), myo-inositol 1-phosphate synthase and myo-inositol
CC monophosphatase-3 (see AAW97883), all enzymes involved in phytate
CC metabolism, are claimed. The invention relates to the use of such genes
CC to reduce the levels of phytate, and/or increase the levels of non-
CC phytate phosphorus, in plants used for food or feed. The genes are
CC especially used to improve the nutritional content of plants such as corn
CC and soybean. Transgenic plants, and seed produced by them, are claimed
XX
SQ Sequence 510 AA;

Query Match 14.9%; Score 76; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 7.5e-67;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	369	SNAILTEPGEHPDHVVVVKVVPYVGDSPKRAMDEYTSEIFMGKNTIVLHNTCEDSLAAP	428
DB	369	SNAILTEPGEHPDHVVVVKVVPYVGDSPKRAMDEYTSEIFMGKNTIVLHNTCEDSLAAP	428
QY	429	IILDVTLAELSTRIQ	444
DB	429	IILDVTLAELSTRIQ	444

Search completed: June 7, 2005, 17:06:33
Job time : 94 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:58:47 ; Search time 24.5 Seconds
(without alignments)
2002.880 Million cell updates/sec

Title: US-10-718-952-12
Perfect score: 510
Sequence: 1 MEIENFKVSEPNVXKTEETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	14.9	510	2	T01647 inositol-3-phospha
2	67	13.1	507	2	S52648 inositol-3-phospha
3	67	13.1	509	2	T08436 inositol-3-phospha
4	67	13.1	510	2	T04399 inositol-3-phospha
5	67	13.1	510	2	T50021 inositol-3-phospha
6	67	13.1	510	2	S60302 inositol-3-phospha
7	67	13.1	510	2	D84610 probable myo-inos
8	67	13.1	511	2	T05017 inositol-3-phospha
9	67	13.1	512	2	T12438 inositol-3-phospha
10	48	9.4	511	2	T10964 inositol-3-phospha
11	15	2.9	430	2	T46317 hypothetical prote
12	13	2.5	525	2	T18569 inositol-3-phospha
13	13	2.5	555	2	A30902 inositol-3-phospha
14	12	2.4	520	1	S45452 inositol-3-phospha
15	8	1.6	210	1	B48357 minor outer capsid
16	8	1.6	210	2	S25544 nonstructural prot
17	8	1.6	378	2	F95844 probable dihydrol
18	8	1.6	973	2	AB2340 hypothetical prote
19	7	1.4	42	1	I65746 tropomyosin - huma
20	7	1.4	96	1	WMVZP3 P3 protein - fowlp
21	7	1.4	97	2	JQ0527 capsid assembly pr
22	7	1.4	105	2	B82149 conserved hypothet
23	7	1.4	125	2	AB1995 hypothetical prote
24	7	1.4	147	2	T02273 hypothetical prote
25	7	1.4	157	2	S77352 hypothetical prote
26	7	1.4	178	2	S44910 ZK666.1 protein -
27	7	1.4	191	2	AH0431 probable exported
28	7	1.4	197	2	T15106 hypothetical prote
29	7	1.4	201	2	AH1812 hypothetical prote

30	7	1.4	209	2	A71313 probable V-type AT
31	7	1.4	212	2	T35864 probable cyclhexa
32	7	1.4	231	2	B83111 50S ribosomal prot
33	7	1.4	236	2	T33333 hypothetical prote
34	7	1.4	280	2	T47572 Machado-Joseph dis
35	7	1.4	282	2	AE2618 rRNA methylase (Im
36	7	1.4	290	2	D97400 rRNA methylase (AP
37	7	1.4	292	2	T30321 hypothetical prote
38	7	1.4	294	2	C82497 ribose ABC transpo
39	7	1.4	294	2	AE2457 hypothetical prote
40	7	1.4	299	2	E84262 rhamnosyl transfer
41	7	1.4	300	2	S16430 hypothetical prote
42	7	1.4	308	2	T08796 tropomyosin - huma
43	7	1.4	309	2	C83017 probable short cha
44	7	1.4	309	2	T02613 hypothetical prote
45	7	1.4	311	2	D64209 hydroxymethylgluta

ALIGNMENTS

RESULT 1
T01647
inositol-3-phosphate synthase (EC 5.5.1.4) - maize
C/Species: Zea mays (maize)
C/Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
C/Accession: T01647
R/Larson, S.R.; Raboy, V.
Submitted to the EMBL Data Library, March 1998
A/Description: Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.
A/Reference number: Z14366
A/Accession: T01647
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-510 <IAR>
A/Cross-references: UNIPROT:O9FPK7; EMBL:AF056326; NID:g3108052; PTDN:AACT15756.1; PID:g31
A/Experimental source: strain Early ACR; leaf
C/Genetics:
A/Gene: INO1
C/Function:
A/Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-ino
A/Pathway: inositol biosynthesis
A/Note: NAD cofactor
C/Keywords: intramolecular lyase; isomerase; NAD

Query Match 14.9%; Score 76; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 7.7e-70;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 SNAIIYEPGEHDPVVIKYVPYVDSKRAMDEYTSIFMGKNTVILNTEGSLAAP 428
|||||
DB 369 SNAIIYEPGEHDPVVIKYVPYVDSKRAMDEYTSIFMGKNTVILNTEGSLAAP 428

QY 429 IILDVLAELSTRIQ 444
|||||
DB 429 IILDVLAELSTRIQ 444

RESULT 2
S52648
inositol-3-phosphate synthase (EC 5.5.1.4) - Citrus paradisi
C/Species: Citrus paradisi
C/Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C/Accession: S52648
R/Holland, D.
Submitted to the EMBL Data Library, April 1994
A/Reference number: S52648
A/Accession: S52648
A/Molecule type: DNA
A/Residues: 1-507 <HOL>
A/Cross-references: UNIPROT:P42802; GB:Z32632; NID:g602564; PTDN:CAA83565.1; PID:g602565
C/Genetics:

QY 301 LIIGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQFRSKEISKS 360
 DB 301 LIIGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQFRSKEISKS 360
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367
 RESULT 7
 D84610
 Probable myo-inositol 1-phosphate synthase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Aug-2004
 C:Accession: D84610
 R:Lin, X.; Kaul, S.; Rounsailey, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.B.; Umayam, L.; Tallon, L.;
 eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: AB4420; MUID:20083487; PMID:10617197
 A:Accession: D84610
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-510 <STO->
 A:Cross-references: UNIPROT:Q38862; GB:AE002093; NID:g4567202; PIDN:AAD23618.1; GSPDB:GN
 A:Gene: At2g22240
 A:Map position: 2
 C:Superfamily: Myo-inositol-1-phosphate synthase
 Query Match 13.1%; Score 67; DB 2; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1.5e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 301 LIIGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQFRSKEISKS 360
 DB 301 LIIGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQFRSKEISKS 360
 QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367
 RESULT 8
 T05017
 Inositol-3-phosphate synthase (EC 5.5.1.4) T19P19.190 [similarity] - Arabidopsis thaliana
 N:Alternate names: protein T19P19.190
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 C:Accession: T05017
 R:Bevan, M.; Monfort, A.; Casacuberta, E.; Puigdomench, P.; Hoheisel, J.; Mewes, H.W.;
 submitted to the Protein Sequence Database, April 1998
 A:Reference number: Z15394
 A:Accession: T05017
 A:Molecule type: DNA
 A:Residues: 1-511 <BEV->
 A:Cross-references: UNIPROT:P42801; EMBL:AL022605
 A:Experimental source: cultivar Columbia; BAC clone T19P19
 C:Genetics:
 A:Map position: 4
 A:Intons: 64/2; 132/3; 215/2; 291/1; 329/3; 388/3; 451/3
 A>Note: T19P19.190
 C:Superfamily: myo-inositol-1-phosphate synthase
 C:Keywords: intramolecular lyase; isomerase; NAD
 Query Match 13.1%; Score 67; DB 2; Length 511;
 Best Local Similarity 100.0%; Pred. No. 1.5e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 301 LIIGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQFRSKEISKS 360
 DB 301 LIIGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQFRSKEISKS 360

DB 302 LIIGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQFRSKEISKS 361
 QY 361 NVVDDMV 367
 DB 362 NVVDDMV 368
 RESULT 9
 T12438
 Inositol-3-phosphate synthase (EC 5.5.1.4) - common ice plant
 C:Species: Mesembryanthemum crystallinum (common ice plant)
 C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
 C:Accession: T12438
 R:Rishitani, M.; Majumder, A.L.; Bornhauser, A.; Michalowski, C.B.; Jensen, R.G.; Bohmert,
 Plant J. 9, 537-548, 1996
 A>Title: Coordinate transcriptional induction of myo-inositol metabolism during environme
 A:Reference number: Z17518; MUID:96208959; PMID:8624516
 A:Accession: T12438
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-512 <ISH->
 A:Cross-references: UNIPROT:Q40271; EMBL:U32511; NID:g975887; PIDN:AAB03687.1; PID:g97588
 C:Superfamily: myo-inositol-1-phosphate synthase
 C:Keywords: intramolecular lyase; isomerase
 Query Match 13.1%; Score 67; DB 2; Length 512;
 Best Local Similarity 100.0%; Pred. No. 1.5e-60;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 301 LIIGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQFRSKEISKS 360
 DB 303 LIIGDDFKSGQTKMSVLVDPLVAGIKRPTISVSYNHLGNNDGMNLSAPQFRSKEISKS 362
 QY 361 NVVDDMV 367
 DB 363 NVVDDMV 369
 RESULT 10
 T10964
 Inositol-3-phosphate synthase (EC 5.5.1.4) - kidney bean
 N:Alternate names: IL-myo-inositol 1-phosphate synthase
 C:Species: Phaseolus vulgaris (kidney bean)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C:Accession: T10964
 R:Wang, X.; Johnson, M.D.
 submitted to the EMBL Data Library, October 1995
 A:Reference number: Z17234
 A:Accession: T10964
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-511 <MAN->
 A:Cross-references: UNIPROT:Q41107; EMBL:U38920; NID:g1066282; PID:g1066283
 A:Experimental source: serran Taylor's horticultural; root
 C:Function:
 A:Description: catalyzes reversible conversion of D-glucose 6-phosphate to 1L-myo-inosit
 A:Pathway: myo-inositol biosynthesis
 C:Superfamily: myo-inositol-1-phosphate synthase
 C:Keywords: intramolecular lyase; isomerase; NAD
 Query Match 9.4%; Score 48; DB 2; Length 511;
 Best Local Similarity 100.0%; Pred. No. 6.4e-41;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 320 DFLVAGIKRPTISVSYNHLGNNDGMNLSAPQFRSKEISKS NVVDDMV 367
 DB 321 DFLVAGIKRPTISVSYNHLGNNDGMNLSAPQFRSKEISKS NVVDDMV 368
 RESULT 11
 T46317
 Hypothetical protein DKFZp434A0612.1 - human
 C:Species: Homo sapiens (man)

C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C/Accession: T46317
R:Duetschhoff, A.; Iabner, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z23035
A/Accession: T46317
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-430 <AAA>
A/Cross-references: UNIPROT:Q9NSU0; EMBL:AL137749
A/Experimental source: adult testis; clone DKFZP434A0612
C/Genetics:
A/Note: DKFZP434A0612.1
C/Superfamily: myo-inositol-1-phosphate synthase

Query Match 2.9%; Score 15; DB 2; Length 430;
Best Local Similarity 100.0%; Pred.No. 6.7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 384 VVKKVVPYGD SKRA 398
Db 253 VVKKVVPYGD SKRA 267

RESULT 12
T18569
inositol-3-phosphate synthase (EC 5.5.1.4) - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T18569; T20002
R:Ainacough, R.
submitted to the EMBL Data Library, December 1998
A/Reference number: Z18979
A/Accession: T18569
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-525 <WIL>
A/Cross-references: UNIPROT:Q18664; EMBL:AL033535; PIDN:CAA22132.1; CESP:VF13D12L.1
A/Experimental source: clone VF13D12L
R:Gajdarsky, S.
submitted to the EMBL Data Library, March 1996
A/Reference number: Z19209
A/Accession: T20002
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-525 <W12>
A/Cross-references: EMBL:Z69902; PIDN:CAA93771.1; GSPDB:GN00020; CESP:VF13D12L.1
A/Experimental source: clone C47D12
C/Genetics:
A/Map position: 2
A/Introns: 106/2; 287/1; 411/2
C/Superfamily: myo-inositol-1-phosphate synthase
C/Keywords: intramolecular lyase; isomerase

Query Match 2.5%; Score 13; DB 2; Length 525;
Best Local Similarity 100.0%; Pred.No. 9.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 331 SIVSYNHLGNNDG 343
Db 337 SIVSYNHLGNNDG 349

RESULT 13
A30902
inositol-3-phosphate synthase (EC 5.5.1.4) [validated] - yeast (*Saccharomyces cerevisiae*)
N/Alternate names: protein j0610; protein YJL153c
C/Species: *Saccharomyces cerevisiae*
C/Date: 18-Apr-1989 #sequence_revision 08-Sep-1995 #text_change 16-Aug-2004
C/Accession: S55160; B32209; S56935; S71644; A30827; A30902
R:Katsoulou, C.; Tzermia, M.; Alexandraki, D.
submitted to the EMBL Data Library, May 1995

A/Description: The complete sequence of a 40.7 kb segment located on the left arm of yeast
Yeast Hypothetical proteins.
A/Reference number: S55159
A/Accession: S55160
A/Molecule type: DNA
A/Residues: 1-555 <KAT>
A/Cross-references: EMBL:X87371; NID:G854542; PID:G854544
R:Dean-Johnson, M.; Henry, S.A.
J. Biol. Chem. 264, 1274-1283, 1989
A/Title: Biosynthesis of inositol in yeast. Primary structure of myo-inositol-1-phosphate
A/Reference number: A32209; MUID:85093118; PMID:2642902
A/Accession: B32209
A/Molecule type: DNA
A/Residues: 23-35, 'RL', 37-81, 'FE', 83-87, 'TRNYAHVWR', 88, 'QQW', 92-103, 'WPRYWRISTWMS', 116-1
VISFQISPSFSAVL' <DEA>
A/Cross-references: EMBL:J04453
R:Katsoulou, C.; Tzermia, M.; Alexandraki, D.
submitted to the Protein Sequence Database, September 1995
A/Reference number: S56912
A/Accession: S56935
A/Molecule type: DNA
A/Residues: 1-555 <KAW>
A/Cross-references: EMBL:Z49428; NID:G1015570; PID:G1015571; MIPS:YJL153C
R:Katsoulou, C.; Tzermia, M.; Tavernarakis, N.; Alexandraki, D.
Yeast 12, 787-797, 1996
A/Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome X
of chromosome XI.
A/Reference number: S71643; MUID:96408771; PMID:8813765
A/Accession: S71644
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-555 <KAF>
A/Cross-references: EMBL:X87371; NID:G854542; PIDN:CAA60802.1; PID:G854544
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
C/Genetics:
A/Genes: SGD:INO1
A/Cross-references: SGD:S0003689; MIPS:YJL153C
A/Map position: 10L
C/Complex: homotetramer
C/Function:
A/Description: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo-
A/Pathway: inositol biosynthesis
A/Note: requires NAD
C/Superfamily: Myo-inositol-1-phosphate synthase
C/Keywords: cytosol; homotetramer; intramolecular lyase; isomerase; NAD

Query Match 2.5%; Score 13; DB 2; Length 555;
Best Local Similarity 100.0%; Pred.No. 9.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 INGSFQNTFVPEL 291
Db 315 INGSFQNTFVPEL 327

RESULT 14
S45452
inositol-3-phosphate synthase (EC 5.5.1.4) - yeast (*Candida albicans*)
C/Species: *Candida albicans*
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S45452
R:KL19, L.S.; Zobel, P.A.; Devry, C.G.; Losberger, C.
Yeast 10, 789-800, 1994
A/Title: Comparison of INO1 gene sequences and products in *Candida albicans* and *Saccharon*
A/Reference number: S45452; MUID:95066381; PMID:7975896
A/Accession: S45452
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-520 <KLI>
A/Cross-references: UNIPROT:P42800; EMBL:L22737; NID:G413758; PIDN:AAA62849.1; PID:G69575
C/Genetics:
A/Genes: INO1
C/Function:

A:Description: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myc
A>Note: requires NAD
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 2.4%; Score 12; DB 2; Length 520;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 279 INGSPOINTFVG 290
Db 285 INGSPOINTFVG 296

RESULT 15

B48357
minor outer capsid protein - porcine rotavirus C (strain Cowden)
N:Alternate names: nonstructural protein NS26
C:Species: porcine rotavirus C
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: B48357
R:Bremont, M.; Chabanne-Vaucherot, D.; Cohen, J.
Arch. Virol. 130, 85-92, 1993
A:Title: Sequence analysis of three non structural proteins of a porcine group C (Cowden
A:Reference number: A48357; MUID:93277387; PMID:8389118
A:Accession: B48357
A:Molecule type: genomic RNA
A:Residues: 1-210 <BRE>
A:Cross-references: UNIPROT:P36358
C:Genetics:
A:Map position: segment 10
C:Superfamily: rotavirus minor outer capsid protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:30,120/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.6%; Score 8; DB 1; Length 210;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 313 KMKSVLVD 320
Db 198 KMKSVLVD 205

Search completed: June 7, 2005, 17:10:36
Job time : 24.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:49:04 ; Search time 86 Seconds
(without alignments)
3036.749 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 510

Sequence: 1 MEIENPKVSPVVKYETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Listing first 45 summaries

Database : UniProt 03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	236	46.3	510	2	094C02
2	236	46.3	510	2	094K03
3	32	18.0	472	2	094G22
4	76	14.9	510	1	INO1_MAIZE
5	76	14.9	510	1	INO1_NICPA
6	76	14.9	510	1	INO1_TOBAC
7	76	14.9	512	2	07XZ66
8	73	14.3	510	2	09AV99
9	72	14.1	510	2	0944C3
10	67	13.1	364	2	09AWG8
11	67	13.1	443	2	08H1B7
12	67	13.1	507	1	INO1_CITPA
13	67	13.1	509	2	09AR12
14	67	13.1	510	1	INO1_BRANA
15	67	13.1	510	1	INO1_HORVU
16	67	13.1	510	1	INO1_ORYSA
17	67	13.1	510	1	INO1_ORYSA
18	67	13.1	510	1	INO1_SSPIN
19	67	13.1	510	1	INO1_SPIPO
20	67	13.1	510	1	INO1_ARATH
21	67	13.1	510	1	INO1_ARATH
22	67	13.1	510	2	07XJCO
23	67	13.1	512	1	INO1_ARATH
24	67	13.1	512	1	INO1_MESCR
25	55	10.8	509	2	08SSN2
26	48	9.4	511	1	INO1_PRAVU
27	47	9.2	164	2	09FT40
28	38	7.5	186	2	042548
29	36	7.1	409	2	0846A1
30	29	5.7	565	1	INO1_DROME
31	27	5.3	135	2	09AT03

32	27	5.3	166	2	09FRP2	09fup2 lycopersico
33	23	4.5	150	2	09SU04	09sua4 brachiosco
34	23	4.5	560	2	06DDT1	06ddt1 xenopus lae
35	23	4.5	563	2	07ZXY0	07zxy0 xenopus lae
36	21	4.1	561	2	07P2B9	07p2b9 anopheles g
37	17	3.3	526	2	07Z525	07z525 homo sapien
38	17	3.3	557	2	09JHU9	09jhu9 m myo-inosi
39	17	3.3	558	2	06NXT5	06nxt5 homo sapien
40	17	3.3	558	2	09H2Y2	09h2y2 homo sapien
41	17	3.3	558	2	09NPH2	09nph2 homo sapien
42	17	3.3	558	2	09NPH2	09nph2 homo sapien
43	16	3.1	220	2	06UP00	06up00 aspergillus
44	16	3.1	537	2	087IUS	087ius neurospora
45	15	2.9	181	2	09BT65	09bt65 homo sapien

ALIGNMENTS

RESULT 1					
ID	094C02	PRELIMINARY;	PRT;	510 AA.	
AC	094C02;				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)				
DE	Myo-inositol-1-phosphate synthase (EC 5.5.1.4).				
OS	Glycine max (Soybean).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.				
OX	NCBI_TaxID=3847;				
RN	(1)				
RA	SEQUENCE FROM N.A.				
RP	Hitc W.D., Carlson T.J., Kerr P., Sebastian S.;				
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AY038802; AKK72098.1; -				
DR	HSSP; P12986; IPIH.				
DR	GO; GO:004512; F:inositol-3-phosphate synthase activity; IEA.				
DR	GO; GO:0016853; F:isomerase activity; IEA.				
DR	GO; GO:006021; P:myo-inositol biosynthesis; IEA.				
DR	GO; GO:008654; P:phospholipid biosynthesis; IEA.				
DR	InterPro: IPR002587; Inos-1-P_synth.				
DR	Pfam: PF01658; Inos-1-P_synth; 1.				
KW	Isomerase.				
SQ	SEQUENCE 510 AA; 56475 MW; A1E4C7F1643918E CRC64;				
Query Match					
Best Local Similarity 46.3%; Score 236; DB 2; Length 510;					
Matches 236; Similarity 100.0%; Pred. No. 4.8e-238; Indels 0; Gaps 0;					
Mismatch 0; Mismatches 0; Indels 0; Gaps 0;					
QY	176 PLPGIYDDPFIANOEERANNVYKGTKEQVOQIIKDIKAFKEATKVDKVVVLMANTER	235			
DB	176 PLPGIYDDPFIANOEERANNVYKGTKEQVOQIIKDIKAFKEATKVDKVVVLMANTER	235			
QY	236 YSNLVVGLNDTMENTLLAAVDRNEAETSPSTLYAIAACWENVPFINGSPONTVPGLIDA	295			
DB	236 YSNLVVGLNDTMENTLLAAVDRNEAETSPSTLYAIAACWENVPFINGSPONTVPGLIDA	295			
QY	296 IARNTLIGGDDPKSGOTKMSVLDPLVGAGIKPTISIVSYNHLGNNDGNTLSAPOTFRSK	355			
DB	296 IARNTLIGGDDPKSGOTKMSVLDPLVGAGIKPTISIVSYNHLGNNDGNTLSAPOTFRSK	355			
QY	356 EISKSNVVDVNNVSNALIEPEGEHPDHVVVVKVPPYVGSKRAMEYTSSEIFMGK 411				
DB	356 EISKSNVVDVNNVSNALIEPEGEHPDHVVVVKVPPYVGSKRAMEYTSSEIFMGK 411				
RESULT 2					
ID	094K03	PRELIMINARY;	PRT;	510 AA.	
AC	094K03;				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				

```
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DR Myo-inositol-3-phosphate synthase.
CN Name=MPS;
OS Glycine max (soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21196082; PubMed=11299373; DOI=10.1104/pp.125.4.1941;
RA Hegeman C.E., Good L.L., Grabau E.A.;
RT "Expression of D-myo-inositol-3-phosphate synthase in soybean.
RT Implications for phylic acid biosynthesis."
RL Plant Physiol. 125:1941-1948(2001).
DR EMBL; AF293970; AAK4986.1; -.
DR HSBP; P11986; 1PIH.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
DR NON_TER 472 472
SQ SEQUENCE 472 AA; 52296 MW; 237204E1A370560F CRC64;
```

```
Query Match 10.0%; Score 92; DB 2; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGDDDFSGGQTKKSVLVDPLVAGIKPTSTIVYNHGNNDGNLAPOTFRSKETS 360
DB 301 LIGDDDFSGGQTKKSVLVDPLVAGIKPTSTIVYNHGNNDGNLAPOTFRSKETS 360
QY 361 NVVDDMVNSNAIILYEPGEHPDHVVVYIKVPYV 392
DB 361 NVVDDMVNSNAIILYEPGEHPDHVVVYIKVPYV 392

RESULT 4
INOI_MAIZE STANDARD; PRT; 510 AA.
AC Q9PRK7; O65196;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
DE synthase) (Mt-1-P synthase) (IPS).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Early ACR; TISSUE=leaf;
RA Larson S.R., Raboy V.;
RT "Linkage mapping maize and barley myo-inositol 1-phosphate synthase
RT genes."
RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Shukla S., Vantolai T.T.;
RT "Genomic sequence of maize myo-inositol 1-phosphate synthase gene."
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
CC phosphate.
CC -1- COFACTOR: NAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF056326; AAC15756.1; -.
DR EMBL; AF323175; AAG40328.1; -.
DR PIR; T01647; T01647.
DR HSBP; P11986; 1PIH.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
FW CONFLICT 18 M -> T (in Ref. 2).
FT CONFLICT 351 A -> T (in Ref. 2).
SQ SEQUENCE 510 AA; 56245 MW; DA59DEC391CEB6D CRC64;

Query Match 14.9%; Score 76; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 3.1e-70;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 SNAIILYEPGEHPDHVVVYIKVPYVGDSCRAMDEYSEIFMGSKNTIVLHNTCESSLAP 428
DB 369 SNAIILYEPGEHPDHVVVYIKVPYVGDSCRAMDEYSEIFMGSKNTIVLHNTCESSLAP 428
```

QY 429 ILLDLVLAELSTRIO 444
 DB 429 ILLDLVLAELSTRIO 444

RESULT 5
 INOI_NICPA STANDARD; PRT; 510 AA.

AC 09SSV4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (Mt-1-P synthase) (IPS).
 DE Name=INPS1;
 OS Nicotiana paniculata.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 CC Lamiales; Solanales; Solanaceae; Nicotiana.
 CC NCBI_TaxId=62141;
 RN RP
 RA Hashimoto A., Yamada S., Komori T.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB032073; BAA84084.1; -.
 DR HSSP; P11986; 1PIH.
 DR InterPro: IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 KW SEQUENCE 510 AA; 56385 MW; 415B81C27A267666 CRC64;

Query Match 14.9%; Score 76; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 3.1e-70;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 SNAIIYEGEHDPHVVIKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLHNTCEDSLAAP 428
 DB 369 SNAIIYEGEHDPHVVIKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLHNTCEDSLAAP 428

QY 429 ILLDLVLAELSTRIO 444
 DB 429 ILLDLVLAELSTRIO 444

RESULT 6
 INOI_TOBAC STANDARD; PRT; 510 AA.

AC 09LW96;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (Mt-1-P synthase) (IPS).
 OS Nicotiana tabacum (Common tobacco).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;

OC Lamiales; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxId=40937;
 RN RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RX MEDLINE=20399434; PubMed=10945337;
 RA Hara K., Yagi M., Koizumi N., Kusano T., Sano H.;
 RT "Screening of wound-responsive genes identifies an immediate-early expressed gene encoding a highly charged protein in mechanically wounded tobacco plants.";
 RL Plant Cell Physiol. 41:684-691(2000).
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-phosphate.
 CC -1- COFACTOR: NAD (By similarity).
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB009881; BAA95788.1; -.
 DR HSSP; P11986; 1PIH.
 DR InterPro: IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 KW SEQUENCE 510 AA; 56369 MW; 4BA8FDDA5DBFED4D CRC64;

Query Match 14.9%; Score 76; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 3.1e-70;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 SNAIIYEGEHDPHVVIKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLHNTCEDSLAAP 428
 DB 369 SNAIIYEGEHDPHVVIKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLHNTCEDSLAAP 428

QY 429 ILLDLVLAELSTRIO 444
 DB 429 ILLDLVLAELSTRIO 444

RESULT 7
 QYXZB6 PRELIMINARY; PRT; 512 AA.

AC 07XZB6;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Inositol 1-phosphate synthase (EC 5.5.1.4).
 GN Name=PINOI;
 OS Porteresia coarctata.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrharioideae; Oryzae; Porteresia.
 CC NCBI_TaxId=77588;
 RN RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22912548; PubMed=14550537; DOI=10.1016/S0014-5793(03)00974-8;
 RA Majumder A.L., Chatterjee A., Ghosh Dasgupta K., Majee M.;
 RT "Diversification and evolution of L-myo-inositol 1-phosphate synthase.";
 RL FEBS Lett. 553:3-10(2003).
 RN RN
 RP SEQUENCE FROM N.A.
 RA Majumder A.N.L., Majee M.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RA Majumder A.N.L., Majee M.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF412340; AAF74579.1; --
 DR HSSP; P11986; 1PKF.
 DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR002587; Inos-1_P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 KW Isomerase.
 SQ SEQUENCE 512 AA; 56651 MW; 69402D547EE1D7F6 CRC64;

Query Match 14.9%; Score 76; DB 2; Length 512;
 Best Local Similarity 100.0%; Pred. No. 3.2e-70;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 369 SNAIYEPGEHPDHVVVKYVPYVGDGSKRAMDEYTSIFMGKNTIVLHNTCEDSLAAPTLL 428
 DB 371 SNAIYEPGEHPDHVVVKYVPYVGDGSKRAMDEYTSIFMGKNTIVLHNTCEDSLAAPTLL 430
 OY 429 IILDVLLAELSTRIO 444
 DB 431 IILDVLLAELSTRIO 446

RESULT 8

O9AV99 PRELIMINARY; PRT; 510 AA.

ID O9AV99
 AC O9AV99;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Myo-inositol-1-phosphate synthase.
 GN Name=MTP;
 OS Avena sativa (Oat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Avenae; Avena
 OC NCBI_TaxID=4498;
 OK [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Seed;
 RC MEDLINE=99097040; PubMed=9880347; DOI=10.1104/pp.119.1.65;
 RX Yoshida Y.T., Wada T., Koyama H., Mizobuchi R., Naito S.;
 RA "Temporal and spatial patterns of accumulation of the transcript of
 RT Myo-inositol-1-phosphate synthase and phytin-containing particles
 during seed development in rice."
 RL Plant Physiol. 119:65-72(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Seed;
 RA Toyonaga D., Saneoka H.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB059557; BAB40956.2; --
 DR HSSP; P11986; 1PK.
 DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
 DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 SQ SEQUENCE 510 AA; 56130 MW; A0F9D980CA0C6CC CRC64;

Query Match 14.3%; Score 73; DB 2; Length 510;
 Best Local Similarity 100.0%; Pred. No. 4.4e-67;
 Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 372 ILYEPGEHPDHVVVKYVPYVGDGSKRAMDEYTSIFMGKNTIVLHNTCEDSLAAPTLL 431
 DB 372 ILYEPGEHPDHVVVKYVPYVGDGSKRAMDEYTSIFMGKNTIVLHNTCEDSLAAPTLL 431
 OY 432 DVLVLAELSTRIO 444

DB 432 DVLVLAELSTRIO 444

RESULT 9

ID O944C3 PRELIMINARY; PRT; 510 AA.

AC O944C3
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Myo-inositol-1-phosphate synthase.
 GN Name=INPS;
 OS Suaeda salsa.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Amaranthaceae; Suaeda.
 OK NCBI_TaxID=126914;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang J.P., Wang P.P., Sun Y.F., Zhao Y.X., Zhang H.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF438793; AAL28131.1; --
 DR HSSP; P11986; 1PK.
 DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
 DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 SQ SEQUENCE 510 AA; 56665 MW; 6C31006D2E1C508A CRC64;

Query Match 14.1%; Score 72; DB 2; Length 510;
 Best Local Similarity 100.0%; Pred. No. 4.9e-66;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 372 ILYEPGEHPDHVVVKYVPYVGDGSKRAMDEYTSIFMGKNTIVLHNTCEDSLAAPTLL 431
 DB 373 ILYEPGEHPDHVVVKYVPYVGDGSKRAMDEYTSIFMGKNTIVLHNTCEDSLAAPTLL 432
 OY 432 DVLVLAELSTRIO 443
 DB 433 DVLVLAELSTRIO 444

RESULT 10

ID O9AWG8 PRELIMINARY; PRT; 364 AA.

AC O9AWG8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Myo-inositol-1-phosphate synthase.
 OS Actinidia arguta.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Ericales; Actinidiaceae; Actinidia.
 OK NCBI_TaxID=64478;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Young mature leaf;
 RA Klages K., Fitzgerald A., Moodie M.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY005128; AAF97409.1; --
 DR HSSP; P11986; 1PK.
 DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
 DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
 DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 SQ SEQUENCE 364 AA; 40256 MW; 1CB64D6FF78127D CRC64;

Query Match 13.1%; Score 67; DB 2; Length 364;
 Best Local Similarity 100.0%; Pred. No. 6.4e-61;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSIIVSYNHLGNDGMNLAPQTRSKSEIKS 360
 DB 155 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSIIVSYNHLGNDGMNLAPQTRSKSEIKS 214

QY 361 NVVDDMV 367
 DB 215 NVVDDMV 221

RESULT 11
 08H1B7 PRELIMINARY; PRT; 443 AA.

AC 08H1B7
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Myo-inositol phosphate synthase (Fragment).
 OS Lolium perenne (Perennial ryegrass).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Poaceae; Lolium.
 OC NCB1_TaxID=4522;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Amiard V., Prud'homme M.-P., Le Dantec C.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY154382; JAKN52772.1; --
 DR HSSP; P11986; 1PKF.
 DR GO; GO:0004512; P:inositol-3-phosphate synthase activity; IEA.
 DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
 DR GO; GO:0006024; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 DR NON TER
 FT 443 443
 SQ SEQUENCE 443 AA; 48572 MW; E8BB1A03779E49FA CRC64;

Query Match 13.1%; Score 67; DB 2; Length 443;
 Best Local Similarity 100.0%; Pred. No. 7.6e-61;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSIIVSYNHLGNDGMNLAPQTRSKSEIKS 360
 DB 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSIIVSYNHLGNDGMNLAPQTRSKSEIKS 360

QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367

RESULT 12
 ID INOI CITPA STANDARD; PRT; 507 AA.

AC P42802;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
 synthase) (MI-1-P synthase) (IPS).
 OS Citrus paradisi (Grapefruit).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Sapindales; Rutaceae; Citrus.
 OC NCB1_TaxID=37656;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tissue-leaf;
 RC MEDLINE=95148748; PubMed=7846170; DOI=10.1104/pp.106.4.1689;
 RA Abu-Abied M., Holland D.;
 RT "The gene c-inol from Citrus paradisi is highly homologous to curl and
 RT inl from Yeast and Spirodela encoding for myo-inositol phosphate
 RT synthase.";

RL Plant Physiol. 106:1689-1689 (1994).
 CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
 CC phosphate.
 CC -1- COFACTOR: NAD.
 CC -1- PATHWAY: Inositol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z32632; CAA83565.1; --
 DR PIR; S52648; S52648.
 DR HSSP; P11986; 1PKF.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 KW SEQUENCE 507 AA; 56334 MW; 45D7892891BDFD6 CRC64;

Query Match 13.1%; Score 67; DB 1; Length 507;
 Best Local Similarity 100.0%; Pred. No. 8.6e-61;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSIIVSYNHLGNDGMNLAPQTRSKSEIKS 360
 DB 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSIIVSYNHLGNDGMNLAPQTRSKSEIKS 360

QY 361 NVVDDMV 367
 DB 361 NVVDDMV 367

RESULT 13
 ID 09AR12 PRELIMINARY; PRT; 509 AA.

AC 09AR12;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Myo-inositol 1-phosphate synthase.
 OS Avicennia marina (Grey mangrove).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Lamiales; Acanthaceae; Acanthaceae incertae sedis; Avicennia.
 OC NCB1_TaxID=82927;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jithesh M.N., Parani M., Parida A.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY028259; AAK21969.1; --
 DR HSSP; P11986; 1PKF.
 DR GO; GO:0004512; P:inositol-3-phosphate synthase activity; IEA.
 DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
 DR GO; GO:0006024; P:phospholipid biosynthesis; IEA.
 DR InterPro; IPR002587; Inos-1-P_synth.
 DR Pfam; PF01658; Inos-1-P_synth; 1.
 SQ SEQUENCE 509 AA; 55978 MW; 23C8D354BAF3BD0F CRC64;

Query Match 13.1%; Score 67; DB 2; Length 509;
 Best Local Similarity 100.0%; Pred. No. 8.6e-61;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSIIVSYNHLGNDGMNLAPQTRSKSEIKS 360
 DB 300 LIGGDDFKSGQTKMSVLVDPLVAGIKPTSIIVSYNHLGNDGMNLAPQTRSKSEIKS 359

QY 361 NVVDDMV 367

Db 360 NVVDDMV 366

RESULT 14

```

INOL_BRNA
ID_INOL_BRNA STANDARD; PRT; 510 AA.
AC 096348;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
synthase) (MI-1-P synthase) (IPS).
OS Brassica napus (rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucoside II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN
RP SEQUENCE FROM N.A.
RA Hussain A., Bourgeois J., Polvi S., Teang E., Keller W.A., Georges F.;
RT "Cloning of a full length cDNA encoding myo-inositol 1-phosphate
synthase from Brassica napus.".
RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = ID-myo-inositol 3-
phosphate.
CC -1- COFACTOR: NAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U66307; AAB06756.2; -.
DR PIR; T08436; T08436.
DR HSSP; P11986; P11986.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56377 MW; A40EB6558D80739 CRC64;

Query Match 13.1%; Score 67; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 8.7e-61;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDFKSGQTKMKSIVLDFVGAGIKPTSIIVSYNHLGNNDGMNLSAPQTRFSKEISK 360
DB 301 LIGGDFKSGQTKMKSIVLDFVGAGIKPTSIIVSYNHLGNNDGMNLSAPQTRFSKEISK 360

QY 361 NVVDDMV 367
DB 361 NVVDDMV 367

RESULT 15
INOL_HORVU
ID_INOL_HORVU STANDARD; PRT; 510 AA.
AC 065195;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
synthase) (MI-1-P synthase) (IPS).
OS Hordeum vulgare (barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;

```

```

OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN
RP SEQUENCE FROM N.A.
RA STRAIN-cv. Harrington;
RA Larson S.R., Raboy V.;
RT "Linkage mapping maize and barley myo-inositol 1-phosphate synthase
genes".
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = ID-myo-inositol 3-
phosphate.
CC -1- COFACTOR: NAD (by similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF056325; AAC17133.1; -.
DR PIR; T04399; T04399.
DR HSSP; P11986; P11986.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56173 MW; EA63138121692724 CRC64;

Query Match 13.1%; Score 67; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 8.7e-61;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 LIGGDFKSGQTKMKSIVLDFVGAGIKPTSIIVSYNHLGNNDGMNLSAPQTRFSKEISK 360
DB 301 LIGGDFKSGQTKMKSIVLDFVGAGIKPTSIIVSYNHLGNNDGMNLSAPQTRFSKEISK 360

QY 361 NVVDDMV 367
DB 361 NVVDDMV 367

```

Search completed: June 7, 2005, 17:09:34
 tod time : 87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 17:00:20 ; Search time 31 Seconds
(without alignments)
1228.098 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 510

Sequence: 1 MFEIHFKEVSPVVKYTEREI.....NIMRACVGLAPENNMLEYK 510

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/pdata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/pdata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/pdata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/pdata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/pdata/1/1aa/PCFUS_COMB.pep:*
- 6: /cgn2_6/pdata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	76	14.9	510	3	US-09-118-442-11 Sequence 11, Appl
2	76	14.9	510	3	US-09-677-064-11 Sequence 11, Appl
3	67	13.1	510	4	US-09-727-628-2 Sequence 2, Appl
4	13	2.5	533	4	US-09-734-237B-73 Sequence 73, Appl
5	13	2.5	534	4	US-09-734-237B-75 Sequence 75, Appl
6	12	2.4	525	4	US-09-248-796A-17234 Sequence 17234, A
7	1481	9	1481	4	US-09-231-899-70 Sequence 70, Appl
8	7	1.4	55	4	US-09-640-211A-2116 Sequence 2116, Ap
9	7	1.4	93	4	US-09-902-540-15884 Sequence 15884, A
10	7	1.4	134	4	US-09-270-767-34609 Sequence 34609, A
11	7	1.4	134	4	US-09-270-767-49826 Sequence 49826, A
12	7	1.4	160	4	US-09-957-641A-19 Sequence 19, Appl
13	7	1.4	191	4	US-09-252-991A-24024 Sequence 24024, A
14	7	1.4	224	2	US-08-272-255-16 Sequence 16, Appl
15	7	1.4	224	5	PCT-US95-08565-16 Sequence 16, Appl
16	7	1.4	234	4	US-09-252-991A-21339 Sequence 21339, A
17	7	1.4	269	4	US-09-902-540-16003 Sequence 16003, A
18	7	1.4	273	3	US-08-235-836C-142 Sequence 142, App
19	7	1.4	273	3	US-08-235-836C-144 Sequence 144, App
20	7	1.4	284	4	US-09-914-259-62 Sequence 62, Appl
21	7	1.4	284	4	US-10-164-595-32 Sequence 32, Appl
22	7	1.4	295	4	US-09-270-767-57050 Sequence 57050, A
23	7	1.4	295	4	US-09-270-767-57401 Sequence 57401, A
24	7	1.4	306	4	US-09-107-532A-5995 Sequence 5995, Ap
25	7	1.4	308	4	US-09-489-039A-11598 Sequence 11598, A
26	7	1.4	322	4	US-09-565-501A-112 Sequence 112, App
27	7	1.4	322	4	US-09-639-206A-112 Sequence 112, App

28	7	1.4	322	4	US-09-874-923-112	Sequence 112, App
29	7	1.4	340	4	US-09-489-039A-9804	Sequence 9804, Ap
30	7	1.4	360	4	US-09-489-039A-9936	Sequence 9936, Ap
31	7	1.4	361	4	US-09-710-279-160	Sequence 160, App
32	7	1.4	364	4	US-09-902-540-16435	Sequence 16435, A
33	7	1.4	381	4	US-09-254-776B-53	Sequence 53, Appl
34	7	1.4	398	4	US-09-949-016-7631	Sequence 7631, Ap
35	7	1.4	399	4	US-09-543-681A-5325	Sequence 5325, Ap
36	7	1.4	443	4	US-09-270-767-46616	Sequence 46616, A
37	7	1.4	443	4	US-09-949-016-11613	Sequence 11613, A
38	7	1.4	446	4	US-09-252-991A-22844	Sequence 22844, A
39	7	1.4	467	4	US-09-489-039A-12726	Sequence 12726, A
40	7	1.4	472	4	US-09-489-039A-8121	Sequence 8121, Ap
41	7	1.4	472	4	US-09-134-000C-3893	Sequence 3893, Ap
42	7	1.4	548	4	US-09-902-540-11870	Sequence 11870, A
43	7	1.4	557	3	US-08-979-608A-5	Sequence 5, Appl
44	7	1.4	557	4	US-09-517-849-5	Sequence 5, Appl
45	7	1.4	557	4	US-09-616-289-5	Sequence 5, Appl

ALIGNMENTS

```
RESULT 1
US-09-118-442-11
; Sequence 11, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-118-442-11
;
;
Query Match          14.9% Score 76; DB 3; Length 510;
Best Local Similarity 100.0%; Pred. No. 4,4e-70;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 369 SNAIYBGEHDPHVVIKYPYVGDSCRANDEYTSKIFMGCKNTIVLHNTCEDSLAAP 428
DB 369 SNAIYBGEHDPHVVIKYPYVGDSCRANDEYTSKIFMGCKNTIVLHNTCEDSLAAP 428
QY 429 IILDVLAELSTRIO 444
DB 429 IILDVLAELSTRIO 444
RESULT 2
US-09-677-064-11
; Sequence 11, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
```

```

; TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-677-064-11
```

```
Query Match          14.9%; Score 76; DB 3; Length 510;
Best Local Similarity 100.0%; Pred. No. 4.4e-70;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 369 SNAIYEPGEHPDHVVIVIKYVPYVGDSSKRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAP 428
DB 369 SNAIYEPGEHPDHVVIVIKYVPYVGDSSKRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAP 428
QY 429 IILDVLVLAELSTRIQ 444
DB 429 IILDVLVLAELSTRIQ 444
```

```

RESULT 3
US-09-727-628-2
; Sequence 2, Application US/09727628
; Patent No. 6791013
; GENERAL INFORMATION:
; APPLICANT: Armstrong, Katherine
; APPLICANT: Hey, Timothy D
; APPLICANT: Folkerts, Otto
; APPLICANT: Smith, Kelley A
; APPLICANT: Hopkins, Nicole L
; TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
; FILE REFERENCE: 50597
; CURRENT APPLICATION NUMBER: US/09/727,628
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/168,612
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-727-628-2
```

```
Query Match          13.1%; Score 67; DB 4; Length 510;
Best Local Similarity 100.0%; Pred. No. 9.7e-61;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 301 LIGGDDFRSGGOTKMSVLVDFVGAGIKPTISVYNHNGNNDGMNLSAPQFRSKETISKS 360
DB 301 LIGGDDFRSGGOTKMSVLVDFVGAGIKPTISVYNHNGNNDGMNLSAPQFRSKETISKS 360
QY 361 NVVDDMV 367
DB 361 NVVDDMV 367
```

```
RESULT 4
US-09-734-237B-73
```

```

; Sequence 73, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-73
```

```
Query Match          2.5%; Score 13; DB 4; Length 533;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 279 INGSPONTFVPGL 291
DB 293 INGSPONTFVPGL 305
```

```

RESULT 5
US-09-734-237B-75
; Sequence 75, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic protein derived from Saccharomyces cerevisiae myo-inosi
; OTHER INFORMATION: tol-1-phosphate synthase, having a glycine residue inserted after
US-09-734-237B-75
```

```
Query Match          2.5%; Score 13; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 279 INGSPONTFVPGL 291
DB 294 INGSPONTFVPGL 306
```

```

RESULT 6
US-09-248-796A-17234
; Sequence 17234, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
```

```

; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17234
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17234

```

```

Query Match 2.4%; Score 12; DB 4; Length 525;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 279 INGSPTFTVPG 290
Db 290 INGSPTFTVPG 301

```

```

RESULT 7
US-09-231-899-70
; Sequence 70, Application US/09231899
; Patent No. 6566583
; GENERAL INFORMATION:
; APPLICANT: Lasner, Michael
; APPLICANT: Metz, James G
; APPLICANT: Facclotti, Daniel
; TITLE OF INVENTION: SCHIZOCHYTRIUM PKS GENES
; FILE REFERENCE: CGNE.131.02US
; CURRENT APPLICATION NUMBER: US/09/231,899
; EARLIER FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/048,650
; EARLIER FILING DATE: 1997-06-04
; EARLIER APPLICATION NUMBER: 09/030,793
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 70
; LENGTH: 1481
; TYPE: PRT
; ORGANISM: Schizochytrium aggregatum
US-09-231-899-70

```

```

Query Match 1.8%; Score 9; DB 4; Length 1481;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 248 ENLLAAVDR 256
Db 1092 ENLLAAVDR 1100

```

```

RESULT 8
US-09-640-211A-2116
; Sequence 2116, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2116
; LENGTH: 55

```

```

; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-640-211A-2116

```

```

Query Match 1.4%; Score 7; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 485 QRAMLEN 491
Db 36 QRAMLEN 42

```

```

RESULT 9
US-09-902-540-15484
; Sequence 15484, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(1589)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15484
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15484

```

```

Query Match 1.4%; Score 7; DB 4; Length 93;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 473 PPGRPVV 479
Db 54 PPGRPVV 60

```

```

RESULT 10
US-09-270-767-34609
; Sequence 34609, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 34609
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE: OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-34609

```

```

Query Match 1.4%; Score 7; DB 4; Length 134;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 286 TFPVGLI 292
Db 99 TFPVGLI 105

```

```
RESULT 11
US-09-270-767-49826
; Sequence 49826, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49826
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-49826

Query Match
Best Local Similarity 1.4%; Score 7; DB 4; Length 134;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 TPVPGHI 292
Db 99 TPVPGHI 105

RESULT 12
US-09-957-641A-19
; Sequence 19, Application US/09957641A
; Patent No. 6770744
; GENERAL INFORMATION:
; APPLICANT: Loliet, John S.
; TITLE OF INVENTION: MODIFIED FACTOR VIII
; FILE REFERENCE: 75-00 US
; CURRENT APPLICATION NUMBER: US/09/957,641A
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 60/234,047
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/236,460
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Porcine
US-09-957-641A-19

Query Match
Best Local Similarity 1.4%; Score 7; DB 4; Length 160;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 476 TPVVNAL 482
Db 119 TPVVNAL 125

RESULT 13
US-09-252-991A-24024
; Sequence 24024, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
```

```
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24024
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24024

Query Match
Best Local Similarity 1.4%; Score 7; DB 4; Length 191;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 APLVPPG 475
Db 85 APLVPPG 91

RESULT 14
US-08-272-255-16
; Sequence 16, Application US/08272255
; Patent No. 5824859
; GENERAL INFORMATION:
; APPLICANT: Cashmore, Anthony R.
; APPLICANT: Ahmad, Margaret
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5824859r18
; STREET: One Liberty place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/272,255
; FILING DATE: 08-JUL-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: UPN-1795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-272-255-16

Query Match
Best Local Similarity 1.4%; Score 7; DB 2; Length 224;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 EDSLILAA 427
Db 8 EDSLILAA 14

RESULT 15
PCT-US95-08565-16
; Sequence 16, Application PC/TUS9508565
```

GENERAL INFORMATION:

APPLICANT: Cashmore, Anthony R.
 APPLICANT: Ahmad, Margaret
 APPLICANT: Lin, Chenhao
 TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
 TITLE OF INVENTION: Using the Same
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
 STREET: One Liberty Place, 46th floor
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/08565
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/272,255
 FILING DATE: 08-JUL-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Leary Ph.D., Kathryn
 REGISTRATION NUMBER: 36,317
 REFERENCE/DOCKET NUMBER: UPN-1795
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-3100
 TELEFAX: (215) 568-3439
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 224 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-08565-16

Query Match 1.4%; Score 7; DB 5; Length 224;

Best Local Similarity 100.0%; Pred. No. 85;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 421 EDSLAA 427

Db 8 EDSLAA 14

Search completed: June 7, 2005, 17:11:48
 Job time : 32 secs

This Page Blank (uspto)

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	510	100.0	510	14	US-10-025-003-12 Sequence 12, Appl
2	510	100.0	510	16	US-10-718-952-12 Sequence 12, Appl
3	423	82.9	510	14	US-10-025-003-14 Sequence 14, Appl
4	423	82.9	510	14	US-10-025-003-16 Sequence 16, Appl
5	423	82.9	510	16	US-10-718-952-14 Sequence 14, Appl
6	423	82.9	510	16	US-10-718-952-16 Sequence 16, Appl
7	236	46.3	510	14	US-10-025-003-2 Sequence 2, Appl
8	236	46.3	510	14	US-10-025-003-10 Sequence 10, Appl
9	236	46.3	510	16	US-10-718-952-2 Sequence 2, Appl
10	236	46.3	510	16	US-10-718-952-10 Sequence 10, Appl
11	236	46.3	511	15	US-10-424-559-213009 Sequence 213009, Appl
12	220	43.1	510	14	US-10-025-003-6 Sequence 6, Appl

13	206	3.1	510	16	US-10-718-955-6	Sequence 6, Appl
14	220	40.4	431	15	US-10-424-599-15863	Sequence 154663, A
15	121	23.7	124	15	US-10-424-599-213004	Sequence 213004, A
16	113	22.2	220	15	US-10-424-599-259439	Sequence 259439, A
17	103	20.2	211	15	US-10-424-599-165505	Sequence 165505, A
18	102	20.0	510	15	US-10-424-599-154864	Sequence 154864, A
19	76	14.9	458	16	US-10-425-115-231843	Sequence 231843, A
20	76	14.9	510	9	US-09-921-332-11	Sequence 11, Appl
21	76	14.9	510	9	US-09-921-330-11	Sequence 11, Appl
22	76	14.9	510	9	US-09-921-329-11	Sequence 11, Appl
23	76	14.9	510	16	US-10-767-701-46278	Sequence 46278, A
24	76	14.9	510	16	US-10-425-115-231852	Sequence 231852, A
25	76	14.9	510	16	US-10-425-115-306814	Sequence 306814, A
26	76	14.9	510	16	US-10-739-930-7635	Sequence 7635, Appl
27	76	14.9	516	15	US-10-425-114-62568	Sequence 62568, A
28	76	14.9	542	15	US-10-425-114-58674	Sequence 58674, A
29	67	13.1	332	15	US-10-425-114-42702	Sequence 42702, A
30	67	13.1	450	16	US-10-425-115-231845	Sequence 231845, A
31	67	13.1	505	15	US-10-442-011-15	Sequence 15, Appl
32	67	13.1	510	9	US-09-727-628-2	Sequence 2, Appl
33	67	13.1	510	16	US-10-437-968-173252	Sequence 173252, A
34	67	13.1	510	16	US-10-425-115-231850	Sequence 231850, A
35	67	13.1	510	16	US-10-425-115-231857	Sequence 231857, A
36	67	13.1	512	15	US-10-425-114-33621	Sequence 33621, A
37	67	13.1	512	15	US-10-425-114-53231	Sequence 53231, A
38	67	13.1	512	15	US-10-425-114-66216	Sequence 66216, A
39	55	10.8	509	16	US-10-437-963-187768	Sequence 187768, A
40	49	9.6	250	15	US-10-424-599-154864	Sequence 154861, A
41	47	9.2	78	15	US-10-424-599-213008	Sequence 213008, A
42	44	8.6	394	16	US-10-363-829-458	Sequence 458, Appl
43	44	8.6	510	16	US-10-425-115-231853	Sequence 231853, A
44	44	8.6	512	15	US-10-425-114-46916	Sequence 46916, A
45	42	8.2	108	16	US-10-425-115-231844	Sequence 231844, A

ALIGNMENTS

```

01      US-10-025-003-12
02      RESULT 1
03      ; Sequence 12, Application US/10025003
04      ; Publication No. US20030074685A1
05      ; GENERAL INFORMATION:
06      ; APPLICANT: Hitz, William
07      ; APPLICANT: Sebastian, Scott
08      ; APPLICANT: Grace, John
09      ; APPLICANT: Streif, Leon
10      ; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
11      ; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
12      ; FILE REFERENCE: BB-1077-C
13      ; CURRENT APPLICATION NUMBER: US/10/025,003
14      ; CURRENT FILING DATE: 2002-05-07
15      ; PRIOR APPLICATION NUMBER: 08/835,751
16      ; PRIOR FILING DATE: APRIL 8, 1997
17      ; PRIOR APPLICATION NUMBER: PCT/US98/06622
18      ; PRIOR FILING DATE: APRIL 7, 1998
19      ; NUMBER OF SEQ ID NOS: 16
20      ; SOFTWARE: Microsoft Office 97
21      ; SEQ ID NO 12
22      ; LENGTH: 510
23      ; TYPE: PRF
24      ; ORGANISM: Glycine max
25      US-10-025-003-12
26
27      Query Match          100.0%; Score 510; DB 14; Length 510;
28      Best Local Similarity 100.0%; Pred. No. 0;
29      Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
30
31      QY      1 MFLENFVSPNNKYETTELQSVYYNYETTELHVNHNNGTYQTWVRKSNVYNQFKNTTHVP 60
32      Db      1 MFLENFVSPNNKYETTELQSVYYNYETTELHVNHNNGTYQTWVRKSNVYNQFKNTTHVP 60
33      1 MFLENFVSPNNKYETTELQSVYYNYETTELHVNHNNGTYQTWVRKSNVYNQFKNTTHVP 60
34      61 KLGVMVLGWMGANNGSTLTGGVIANREDISWATDKIQANVFGLSLTOASAIRVSGSFQGEH 120

```

```
Db 61 KLGWLVGNGNGSTLIGVLANREDISMAWKDKIQOANYGSLTQSAIRVSGFQGBE 120
Qy 121 IYAPFKSLIPMVPDDIVFGWDISNMNLADAMARAKVFDIDLOKQLRPYMESMPLPGI 180
Db 121 IYAPFKSLIPMVPDDIVFGWDISNMNLADAMARAKVFDIDLOKQLRPYMESMPLPGI 180
Qy 181 YPDPFLAANOEBRANNVIKGTQOEVOQI IKDIAFKKATKVDKVVLTMTANTERYSNLV 240
Db 181 YPDPFLAANOEBRANNVIKGTQOEVOQI IKDIAFKKATKVDKVVLTMTANTERYSNLV 240
Qy 241 VGLNDTMENLLAADRNEAISPSTLYAIACMENVPIFNGSPONTFVGLIDLAIARNT 300
Db 241 VGLNDTMENLLAADRNEAISPSTLYAIACMENVPIFNGSPONTFVGLIDLAIARNT 300
Qy 301 LIGGDFKSGQTKMSVLDVFLVGAGIKPTISIVSNHNGNDGMNLSAPQTFRSKEISKS 360
Db 301 LIGGDFKSGQTKMSVLDVFLVGAGIKPTISIVSNHNGNDGMNLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALIYEPGHPDHVVYIKVYPVGDSCRAMDEYTSIFMGKNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGHPDHVVYIKVYPVGDSCRAMDEYTSIFMGKNTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAELSTRIOFKAENEGKFSHFVATILSYLTAKPLVPPGTPVNV 480
Db 421 EDSLAAPIIIDLVLLAELSTRIOFKAENEGKFSHFVATILSYLTAKPLVPPGTPVNV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 2

```
US-10-718-952-12
; Sequence 12, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-12
```

```
Query Match 100.0%; Score 510; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 181 YPDPFLAANOEBRANNVIKGTQOEVOQI IKDIAFKKATKVDKVVLTMTANTERYSNLV 240
Db 181 YPDPFLAANOEBRANNVIKGTQOEVOQI IKDIAFKKATKVDKVVLTMTANTERYSNLV 240
Qy 241 VGLNDTMENLLAADRNEAISPSTLYAIACMENVPIFNGSPONTFVGLIDLAIARNT 300
Db 241 VGLNDTMENLLAADRNEAISPSTLYAIACMENVPIFNGSPONTFVGLIDLAIARNT 300
Qy 301 LIGGDFKSGQTKMSVLDVFLVGAGIKPTISIVSNHNGNDGMNLSAPQTFRSKEISKS 360
Db 301 LIGGDFKSGQTKMSVLDVFLVGAGIKPTISIVSNHNGNDGMNLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALIYEPGHPDHVVYIKVYPVGDSCRAMDEYTSIFMGKNTIVLHNTC 420
Db 361 NVDDMVNSNALIYEPGHPDHVVYIKVYPVGDSCRAMDEYTSIFMGKNTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAELSTRIOFKAENEGKFSHFVATILSYLTAKPLVPPGTPVNV 480
Db 421 EDSLAAPIIIDLVLLAELSTRIOFKAENEGKFSHFVATILSYLTAKPLVPPGTPVNV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 3

```
US-10-025-003-14
; Sequence 14, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-14
```

```
Query Match 82.9%; Score 423; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```

Db      328 KPTSIYSYHNLGNNDGMSLAPQTFRSKISKSNNVDDVNSNAILYEGEHPDHVVYK 387
      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLLAEISTRIOFKA 447
      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLLAEISTRIOFKA 447
Qy      448 ENEGKFHSHFPAVATILSYLTKAPLVPRGTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
      448 ENEGKFHSHFPAVATILSYLTKAPLVPRGTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
Db      508 EYK 510
      508 EYK 510
      508 EYK 510

```

RESULT 4
US-10-025-003-16
; Sequence 16, Application US/10025003
; Publication No. US20030074665A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025, 003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-16

```

Query Match      82.9%; Score 423; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      88 ISMATKDKIQOANYFGSLTQASAIRVGSFOGEIYAPFKSLPMVNPDDIVFGMDISNM 147
      88 ISMATKDKIQOANYFGSLTQASAIRVGSFOGEIYAPFKSLPMVNPDDIVFGMDISNM 147
Db      148 NLADAMAAKAFVDIDLQKQLRPYMESWPLPGIYDPDPIAANOEBRANNVIKGTQOEQVQ 207
      148 NLADAMAAKAFVDIDLQKQLRPYMESWPLPGIYDPDPIAANOEBRANNVIKGTQOEQVQ 207
Qy      208 QIIKDIKAFKATKVDKVVLTANTERYSNLVGLNDTMENLLAADVREAEISPSTLY 267
      208 QIIKDIKAFKATKVDKVVLTANTERYSNLVGLNDTMENLLAADVREAEISPSTLY 267
Db      208 QIIKDIKAFKATKVDKVVLTANTERYSNLVGLNDTMENLLAADVREAEISPSTLY 267
      208 QIIKDIKAFKATKVDKVVLTANTERYSNLVGLNDTMENLLAADVREAEISPSTLY 267
Qy      266 AIACVMEVVPFINGSPONTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLVDPLVGAGI 327
      266 AIACVMEVVPFINGSPONTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLVDPLVGAGI 327
Db      266 AIACVMEVVPFINGSPONTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLVDPLVGAGI 327
      266 AIACVMEVVPFINGSPONTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLVDPLVGAGI 327
Qy      328 KPTSIYSYHNLGNNDGMSLAPQTFRSKISKSNNVDDVNSNAILYEGEHPDHVVYK 387
      328 KPTSIYSYHNLGNNDGMSLAPQTFRSKISKSNNVDDVNSNAILYEGEHPDHVVYK 387
Db      328 KPTSIYSYHNLGNNDGMSLAPQTFRSKISKSNNVDDVNSNAILYEGEHPDHVVYK 387
      328 KPTSIYSYHNLGNNDGMSLAPQTFRSKISKSNNVDDVNSNAILYEGEHPDHVVYK 387
Qy      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLLAEISTRIOFKA 447
      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLLAEISTRIOFKA 447
Db      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLLAEISTRIOFKA 447
      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLLAEISTRIOFKA 447
Qy      448 ENEGKFHSHFPAVATILSYLTKAPLVPRGTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
      448 ENEGKFHSHFPAVATILSYLTKAPLVPRGTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
Db      448 ENEGKFHSHFPAVATILSYLTKAPLVPRGTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
      448 ENEGKFHSHFPAVATILSYLTKAPLVPRGTPVNNALSKORAMLENIMRACVGLAPENNMTL 507

```

```

Qy      508 EYK 510
      508 EYK 510
      508 EYK 510

```

RESULT 5
US-10-718-952-14

```

; Sequence 14, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718, 952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-14

```

```

Query Match      82.9%; Score 423; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      88 ISMATKDKIQOANYFGSLTQASAIRVGSFOGEIYAPFKSLPMVNPDDIVFGMDISNM 147
      88 ISMATKDKIQOANYFGSLTQASAIRVGSFOGEIYAPFKSLPMVNPDDIVFGMDISNM 147
Db      148 NLADAMAAKAFVDIDLQKQLRPYMESWPLPGIYDPDPIAANOEBRANNVIKGTQOEQVQ 207
      148 NLADAMAAKAFVDIDLQKQLRPYMESWPLPGIYDPDPIAANOEBRANNVIKGTQOEQVQ 207
Qy      208 QIIKDIKAFKATKVDKVVLTANTERYSNLVGLNDTMENLLAADVREAEISPSTLY 267
      208 QIIKDIKAFKATKVDKVVLTANTERYSNLVGLNDTMENLLAADVREAEISPSTLY 267
Db      208 QIIKDIKAFKATKVDKVVLTANTERYSNLVGLNDTMENLLAADVREAEISPSTLY 267
      208 QIIKDIKAFKATKVDKVVLTANTERYSNLVGLNDTMENLLAADVREAEISPSTLY 267
Qy      266 AIACVMEVVPFINGSPONTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLVDPLVGAGI 327
      266 AIACVMEVVPFINGSPONTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLVDPLVGAGI 327
Db      266 AIACVMEVVPFINGSPONTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLVDPLVGAGI 327
      266 AIACVMEVVPFINGSPONTFVPGGLDLAIARNTLIGDDPFSGQTKMSVLVDPLVGAGI 327
Qy      328 KPTSIYSYHNLGNNDGMSLAPQTFRSKISKSNNVDDVNSNAILYEGEHPDHVVYK 387
      328 KPTSIYSYHNLGNNDGMSLAPQTFRSKISKSNNVDDVNSNAILYEGEHPDHVVYK 387
Db      328 KPTSIYSYHNLGNNDGMSLAPQTFRSKISKSNNVDDVNSNAILYEGEHPDHVVYK 387
      328 KPTSIYSYHNLGNNDGMSLAPQTFRSKISKSNNVDDVNSNAILYEGEHPDHVVYK 387
Qy      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLLAEISTRIOFKA 447
      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLLAEISTRIOFKA 447
Db      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLLAEISTRIOFKA 447
      388 YVPYVGDSCRAMDEYTSSEIFMGSKNTIVLHNTCEDSLAAPILDLVLLAEISTRIOFKA 447
Qy      448 ENEGKFHSHFPAVATILSYLTKAPLVPRGTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
      448 ENEGKFHSHFPAVATILSYLTKAPLVPRGTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
Db      448 ENEGKFHSHFPAVATILSYLTKAPLVPRGTPVNNALSKORAMLENIMRACVGLAPENNMTL 507
      448 ENEGKFHSHFPAVATILSYLTKAPLVPRGTPVNNALSKORAMLENIMRACVGLAPENNMTL 507

```

GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/718,952
CURRENT FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO: 16
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-718-952-16

Query Match 82.9%; Score 423; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ISMATKDKIQOANYFGSLTQASAIRVSGSEIYAPFKSLPMVNPDDIVFGMDISM 147
DB 88 ISMATKDKIQOANYFGSLTQASAIRVSGSEIYAPFKSLPMVNPDDIVFGMDISM 147
QY 148 NLADAMARAKVPDIDLOKOLRPYMSVPLPGIYDPDFTAAOBERANNVIKGTQEOVQ 207
DB 148 NLADAMARAKVPDIDLOKOLRPYMSVPLPGIYDPDFTAAOBERANNVIKGTQEOVQ 207
QY 208 QIIKIKAKKERTKVDKVVVLTANTERSNLVGNLDMENLLAADRNEAISPSTLY 267
DB 208 QIIKIKAKKERTKVDKVVVLTANTERSNLVGNLDMENLLAADRNEAISPSTLY 267
QY 268 AIAQWENVPFINGSPQNTFVPGILDALAIARNTLIGDDPFKSGQTKMSVLDVVGAGI 327
DB 268 AIAQWENVPFINGSPQNTFVPGILDALAIARNTLIGDDPFKSGQTKMSVLDVVGAGI 327
QY 328 KPTSIYSVNHLCNNDGMLISAPQTFERSKISKSNVDDMVNSNAILYEPGEHPDHVVVK 387
DB 328 KPTSIYSVNHLCNNDGMLISAPQTFERSKISKSNVDDMVNSNAILYEPGEHPDHVVVK 387
QY 388 YVPYVDSKRADEYTSSEIFMGKNTIVLHNTCEBSLAPILDLVLAELSTRIOFKA 447
DB 388 YVPYVDSKRADEYTSSEIFMGKNTIVLHNTCEBSLAPILDLVLAELSTRIOFKA 447
QY 448 ENEGKFSHPVATILSYLTAKPLVPGTPVYNALSKORAMLENIRACVGLAPENMIL 507
DB 448 ENEGKFSHPVATILSYLTAKPLVPGTPVYNALSKORAMLENIRACVGLAPENMIL 507
QY 508 EYK 510
DB 508 EYK 510

RESULT 7
US-10-025-003-2
Sequence 2, Application US/10025003
Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07

PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO: 2
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-025-003-2

Query Match 46.3%; Score 236; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 7.1e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPDFTAAOBERANNVIKGTQEOVOQIIKDKAFKERTKVDKVVVLTANTER 235
DB 176 PLPGIYDPDFTAAOBERANNVIKGTQEOVOQIIKDKAFKERTKVDKVVVLTANTER 235
QY 236 YSNLVGNDTMENTLLAADRNEAISPSTLYAIAQWENVPFINGSPQNTFVPGILDIA 295
DB 236 YSNLVGNDTMENTLLAADRNEAISPSTLYAIAQWENVPFINGSPQNTFVPGILDIA 295
QY 296 IARNTLIGDDPFKSGQTKMSVLDVVGAGIKPTSIYSVNHLCNNDGMLISAPQTFERSK 355
DB 296 IARNTLIGDDPFKSGQTKMSVLDVVGAGIKPTSIYSVNHLCNNDGMLISAPQTFERSK 355
QY 356 EISKSNVDDMVNSNAILYEPGEHPDHVVVKYVPYVDSKRADEYTSSEIFMGSK 411
DB 356 EISKSNVDDMVNSNAILYEPGEHPDHVVVKYVPYVDSKRADEYTSSEIFMGSK 411

RESULT 8
US-10-025-003-10
Sequence 10, Application US/10025003
Publication No. US20030074685A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streitz, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025,003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/835,751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
PRIOR FILING DATE: APRIL 7, 1998
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO: 10
LENGTH: 510
TYPE: PRT
ORGANISM: Glycine max
US-10-025-003-10

Query Match 46.3%; Score 236; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 7.1e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDPDFTAAOBERANNVIKGTQEOVOQIIKDKAFKERTKVDKVVVLTANTER 235
DB 176 PLPGIYDPDFTAAOBERANNVIKGTQEOVOQIIKDKAFKERTKVDKVVVLTANTER 235
QY 236 YSNLVGNDTMENTLLAADRNEAISPSTLYAIAQWENVPFINGSPQNTFVPGILDIA 295
DB 236 YSNLVGNDTMENTLLAADRNEAISPSTLYAIAQWENVPFINGSPQNTFVPGILDIA 295
QY 296 IARNTLIGDDPFKSGQTKMSVLDVVGAGIKPTSIYSVNHLCNNDGMLISAPQTFERSK 355
DB 296 IARNTLIGDDPFKSGQTKMSVLDVVGAGIKPTSIYSVNHLCNNDGMLISAPQTFERSK 355

Db 296 IARNTLIGDDPFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDGNLAPQTFRSK 355
QY 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEMGSK 411
Db 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEMGSK 411

RESULT 9

US-10-718-952-2
; Sequence 2, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Stett, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-2

Query Match 46.3%; Score 236; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 7.1e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDDPFIANAONEERANNVIKGTKEOVQOIIKDIKAFKFAATKVDKVVVLTANTR 235
Db 176 PLPGIYDDPFIANAONEERANNVIKGTKEOVQOIIKDIKAFKFAATKVDKVVVLTANTR 235
QY 236 YSNLVGLNDTMENTLLAAVDRNEAISPSTLYALACWENVPFINGSPONTFVPGILDLA 295
Db 236 YSNLVGLNDTMENTLLAAVDRNEAISPSTLYALACWENVPFINGSPONTFVPGILDLA 295
QY 236 IARNTLIGDDPFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDGNLAPQTFRSK 355
Db 236 IARNTLIGDDPFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDGNLAPQTFRSK 355
QY 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEMGSK 411
Db 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEMGSK 411

RESULT 10

US-10-718-952-10
; Sequence 10, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Stett, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-10

Query Match 46.3%; Score 236; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 7.1e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDDPFIANAONEERANNVIKGTKEOVQOIIKDIKAFKFAATKVDKVVVLTANTR 235
Db 176 PLPGIYDDPFIANAONEERANNVIKGTKEOVQOIIKDIKAFKFAATKVDKVVVLTANTR 235
QY 236 YSNLVGLNDTMENTLLAAVDRNEAISPSTLYALACWENVPFINGSPONTFVPGILDLA 295
Db 236 YSNLVGLNDTMENTLLAAVDRNEAISPSTLYALACWENVPFINGSPONTFVPGILDLA 295
QY 236 IARNTLIGDDPFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDGNLAPQTFRSK 355
Db 236 IARNTLIGDDPFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDGNLAPQTFRSK 355
QY 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEMGSK 411
Db 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEMGSK 411

RESULT 11

US-10-424-599-213009
; Sequence 213009, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 213009
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Glycine max
; NAME/KEY: unsure
; LOCATION: (1)..(511)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1.pap
US-10-424-599-213009

Query Match 46.3%; Score 236; DB 15; Length 511;
Best Local Similarity 100.0%; Pred. No. 7.2e-225;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 PLPGIYDDPFIANAONEERANNVIKGTKEOVQOIIKDIKAFKFAATKVDKVVVLTANTR 235
Db 176 PLPGIYDDPFIANAONEERANNVIKGTKEOVQOIIKDIKAFKFAATKVDKVVVLTANTR 236
QY 236 YSNLVGLNDTMENTLLAAVDRNEAISPSTLYALACWENVPFINGSPONTFVPGILDLA 295
Db 236 YSNLVGLNDTMENTLLAAVDRNEAISPSTLYALACWENVPFINGSPONTFVPGILDLA 296
QY 236 IARNTLIGDDPFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDGNLAPQTFRSK 355
Db 236 IARNTLIGDDPFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDGNLAPQTFRSK 355
QY 297 IARNTLIGDDPFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDGNLAPQTFRSK 356
Db 297 IARNTLIGDDPFKSGQTKKKSVLVDPLVGAGIKPTISVSYNHLGNNDGNLAPQTFRSK 356
QY 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEMGSK 411
Db 356 EISKSNVVDMDVNSNAIILYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEMGSK 411

Db 357 EISKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIFMGSK 412

RESULT 12

US-10-025-003-6

/ Sequence 6, Application US/10025003

/ Publication No. US20030074685A1

/ GENERAL INFORMATION:

/ APPLICANT: Hiltz, William

/ APPLICANT: Sebastian, Scott

/ APPLICANT: Grace, John

/ APPLICANT: Streitz, Leon

/ TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE

/ FILE REFERENCE: BB-1077-C

/ CURRENT APPLICATION NUMBER: US/10/025, 003

/ CURRENT FILING DATE: 2002-05-07

/ PRIOR APPLICATION NUMBER: 08/835, 751

/ PRIOR FILING DATE: APRIL 8, 1997

/ PRIOR APPLICATION NUMBER: PCT/US98/06822

/ PRIOR FILING DATE: APRIL 7, 1998

/ NUMBER OF SEQ ID NOS: 16

/ SOFTWARE: Microsoft Office 97

/ SEQ ID NO 6

/ LENGTH: 510

/ TYPE: PRT

/ ORGANISM: Glycine max

US-10-025-003-6

Query Match 43.1%; Score 220; DB 14; Length 510;

Best Local Similarity 100.0%; Pred. No. 5.3e-209;

Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 PLPGIYDPDFIAANOEBRANNVIKGTQKQVQOIIKDIKAFKATKYVDKVVVLTANTER 235

Db 176 PLPGIYDPDFIAANOEBRANNVIKGTQKQVQOIIKDIKAFKATKYVDKVVVLTANTER 235

Qy 236 YSNLVGGLNDTMENTLLAANDRNEAEISPSSTLYALACWENVPFINGSPQNTFVPGILDIA 295

Db 236 YSNLVGGLNDTMENTLLAANDRNEAEISPSSTLYALACWENVPFINGSPQNTFVPGILDIA 295

Qy 236 IARNTLIGDDPKSGQTKMKSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSK 355

Db 236 IARNTLIGDDPKSGQTKMKSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSK 355

Qy 296 IARNTLIGDDPKSGQTKMKSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSK 355

Db 296 IARNTLIGDDPKSGQTKMKSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSK 355

Qy 356 EISKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSS 395

Db 356 EISKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSS 395

RESULT 13

US-10-718-952-6

/ Sequence 6, Application US/10718952

/ Publication No. US20040128713A1

/ GENERAL INFORMATION:

/ APPLICANT: Hiltz, William

/ APPLICANT: Sebastian, Scott

/ APPLICANT: Grace, John

/ APPLICANT: Streitz, Leon

/ TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE

/ FILE REFERENCE: BB-1077-C

/ CURRENT APPLICATION NUMBER: US/10/718, 952

/ CURRENT FILING DATE: 2003-11-21

/ PRIOR APPLICATION NUMBER: 08/835, 751

/ PRIOR FILING DATE: APRIL 8, 1997

/ PRIOR APPLICATION NUMBER: PCT/US98/06822

/ PRIOR FILING DATE: APRIL 7, 1998

/ NUMBER OF SEQ ID NOS: 16

/ SOFTWARE: Microsoft Office 97

/ SEQ ID NO 6

/ LENGTH: 510

/ TYPE: PRT

/ ORGANISM: Glycine max

US-10-718-952-6

Query Match 43.1%; Score 220; DB 16; Length 510;

Best Local Similarity 100.0%; Pred. No. 5.3e-209;

Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 176 PLPGIYDPDFIAANOEBRANNVIKGTQKQVQOIIKDIKAFKATKYVDKVVVLTANTER 235

Db 176 PLPGIYDPDFIAANOEBRANNVIKGTQKQVQOIIKDIKAFKATKYVDKVVVLTANTER 235

Qy 236 YSNLVGGLNDTMENTLLAANDRNEAEISPSSTLYALACWENVPFINGSPQNTFVPGILDIA 295

Db 236 YSNLVGGLNDTMENTLLAANDRNEAEISPSSTLYALACWENVPFINGSPQNTFVPGILDIA 295

Qy 296 IARNTLIGDDPKSGQTKMKSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSK 355

Db 296 IARNTLIGDDPKSGQTKMKSVLVDPLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSK 355

Qy 356 EISKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSS 395

Db 356 EISKSNVDDMVNSNALIYEPGEHPDHVVVIKYVPYVGDSS 395

RESULT 14

US-10-424-599-154863

/ Sequence 154863, Application US/10424599

/ Publication No. US20040031072A1

/ GENERAL INFORMATION:

/ APPLICANT: La Rosa, Thomas J

/ APPLICANT: Kovalic, David K

/ APPLICANT: Zhou, Yihua

/ APPLICANT: Cao, Yongwei

/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

/ FILE REFERENCE: 38-21(53223)B

/ CURRENT APPLICATION NUMBER: US/10/424, 599

/ CURRENT FILING DATE: 2003-04-28

/ NUMBER OF SEQ ID NOS: 285684

/ SEQ ID NO 154863

/ LENGTH: 431

/ TYPE: PRT

/ ORGANISM: Glycine max

/ FEATURE:

/ OTHER INFORMATION: Clone ID: PAT_MRT3847_110862C.1.pcp

US-10-424-599-154863

Query Match 40.4%; Score 206; DB 15; Length 431;

Best Local Similarity 99.7%; Pred. No. 3.5e-195;

Matches 306; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 88 ISMATKDKIQOANYFSGSLTQASAIRVGSFQGEETIYAPFKSLIPMVPDDIYFGMDISNM 147

Db 88 ISMATKDKIQOANYFSGSLTQASAIRVGSFQGEETIYAPFKSLIPMVPDDIYFGMDISNM 147

Qy 148 NLADAMARAKVFDIDLOKQRLPYMESVPLPGIYDPDFIAANOEBRANNVIKGTQKQVQ 207

Db 148 NLADAMARAKVFDIDLOKQRLPYMESVPLPGIYDPDFIAANOEBRANNVIKGTQKQVQ 207

Qy 208 QIIKDIKAFKATKYVDKVVVLTANTERYSNLVGLNDTMENTLLAANDRNEAEISPSSTLY 267

Db 208 QIIKDIKAFKATKYVDKVVVLTANTERYSNLVGLNDTMENTLLAANDRNEAEISPSSTLY 267

Qy 268 AIACWENVPFINGSPQNTFVPGILDIAIARNTLIGDDPKSGQTKMKSVLVDPLVGAGI 327

Db 268 AIACWENVPFINGSPQNTFVPGILDIAIARNTLIGDDPKSGQTKMKSVLVDPLVGAGI 327

Qy 328 KPTSIYSYNHLGNDGMNLSAPQTFRSKEISKSNVDDMVNSNALIYEPGEHPDHVVVIK 387

Db 328 KPTSIYSYNHLGNDGMNLSAPQTFRSKEISKSNVDDMVNSNALIYEPGEHPDHVVVIK 387

Qy 388 YVPYVGD 394

Db 388 YVPVYCD 394

RESULT 15

US-10-424-599-213004

/ Sequence 213004, Application US/10424599

/ Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223) B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 213004

LENGTH: 124

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_34368C.1.pep

US-10-424-599-213004

Query Match 23.7%; Score 121; DB 15; Length 124;

Best Local Similarity 100.0%; Pred. No. 2.2e-111;

Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	390	PVVGDSKRAMDEYTSRIFMGKNTIVLHNTCEDSLAPILIDLVLAELSTRIOPKAEN	449
DB	4	PVVGDSKRAMDEYTSRIFMGKNTIVLHNTCEDSLAPILIDLVLAELSTRIOPKAEN	63
QY	450	BGKFSFHFVATILSYLTAKPLVPPTPVVNAISKORAMLENIMRACVGLAPENNMILEY	509
DB	64	BGKFSFHFVATILSYLTAKPLVPPTPVVNAISKORAMLENIMRACVGLAPENNMILEY	123
QY	510	K 510	
DB	124	K 124	

Search completed: June 7, 2005, 17:25:34
Job time : 78.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 15:12:39 ; Search time 86 Seconds
(without alignments) 3036.749 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631

Sequence: 1 MFENPFKESPNVKTETETI.....NIMRACVGLAPENNMLEYK 510

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2593	98.6	510	2 094C02	094C02 glycine max
2	2580	98.1	510	2 094K03	094K03 glycine max
3	2475	94.1	510	1 INO1_TOBAC	091W96 nicotiana t
4	2470	93.9	510	1 INO1_NICPA	096Sv4 nicotiana p
5	2452	93.2	510	1 INO1_SESIN	095FV1 sesamum ind
6	2400	91.2	510	1 INO1_WHEBAT	0947U0 triticum ae
7	2399	91.2	510	2 0944G3	0944G3 sueda sals
8	2395	91.0	512	1 INO1_MESCR	040271 mesembryant
9	2384	90.6	510	1 INO2_ARATH	038862 arabidopsis
10	2376	90.3	510	1 INO1_BRANA	096348 brassica na
11	2376	90.3	510	1 INO3_ARATH	091X12 arabidopsis
12	2373.5	90.2	509	2 09AR12	09AR12 avicennia m
13	2361	89.7	510	2 07XJCO	07XJCO xerophyta v
14	2358	89.6	510	1 INO1_SPIRO	042803 epiriodela p
15	2354	89.5	510	1 INO1_MAIZE	095FK7 zea mays (m
16	2354	89.5	510	2 09AV99	09AV99 avena sativ
17	2349.5	89.3	511	1 INO1_ARATH	042801 arabidopsis
18	2345	89.1	510	1 INO1_ORYSA	064193 oryza sativ
19	2338	88.9	510	1 INO1_HORYU	065135 hordeum vul
20	2337	88.8	472	2 094G22	094G22 phaseolus v
21	2328.5	88.5	511	1 INO1_PITPA	041107 phaseolus v
22	2304.5	87.6	507	1 INO1_CITPA	042802 citrus para
23	2221.5	84.4	509	2 08SSN2	08SSN2 oryza sativ
24	2207	76.3	443	2 08H1B7	08H1B7 lolium pere
25	1919.5	73.0	409	2 08AS41	08AS41 ascer tipo
26	1883	71.6	512	2 07XZB6	07XZB6 porteresia
27	1728	65.7	364	2 09AWG8	09AWG8 actinidia a
28	1722.5	65.5	563	2 07ZXVO	07ZXVO xenopus lae
29	1716.5	65.2	560	2 06DDT1	06DDT1 xenopus lae
30	1616	61.4	558	2 09NPH2	09NPH2 homo sapien
31	1616	61.4	558	2 09NVW7	09NVW7 homo sapien

32	1611	61.2	558	2 09H2Y2	09H2Y2 homo sapien
33	1606	61.0	537	2 087JUS	087JUS neurospora
34	1606	61.0	537	2 09JH09	09JH09 m myo-inosi
35	1603	60.9	558	2 06NXT5	06NXT5 homo sapien
36	1602	60.9	565	1 INO1_DROME	097477 drosophila
37	1586.5	60.3	561	2 07PZB9	07PZB9 anopheles g
38	1566.5	59.5	526	2 07Z525	07Z525 homo sapien
39	1525	58.0	526	2 095PT4	095PT4 leishmania
40	1509	57.4	526	2 09N9B7	09N9B7 leishmania
41	1476	56.1	530	2 06CF53	06CF53 yarrowia li
42	1472	55.9	508	2 P90626	P90626 entamoeba h
43	1469.5	55.9	527	2 000815	000815 leishmania
44	1459.5	55.5	520	2 06B3J5	06B3J5 debaryomyce
45	1450	55.1	525	2 074247	074247 pitichia past

ALIGNMENTS

RESULT 1

ID	094C02	PRELIMINARY;	PRT;	510 AA.
AC	094C02;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE	Myo-inositol-1-phosphate synthase (EC 5.5.1.4).			
OS	Glycine max (Soybean).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosidia I; Fabales; Fabaceae; Papilionoideae; Phaseolaceae; Glycine.			
NCBI	NCBI_TaxID=3847;			
RM	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hitz W.D., Carlson T.J., Kerr P., Sebastian S.;			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY038802; AAK72098.1; --			
DR	HSSP; P1986; 1PIH.			
DR	GO; GO:0004512; P:inositol-3-phosphate synthase activity; IEA.			
DR	GO; GO:0016853; P:isomerase activity; IEA.			
DR	GO; GO:0006021; P:myo-inositol biosynthesis; IEA.			
DR	GO; GO:0008654; P:phospholipid biosynthesis; IEA.			
DR	InterPro; IPR002587; Inos-1-P_synth.			
KW	Pfam; PF01658; Inos-1-P_synth; 1.			
KW	Isomerase.			
SC	SEQUENCE 510 AA; 56475 MW; A1B4C77F1643918E CRC64;			
Query Match	98.6%; Score 2593; DB 2; Length 510;			
Best Local Similarity	98.4%; Pred. No. 2.6e-162;			
Matches	502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;			
QY	1 MFENPFKESPNVKTETETIQSIVVNYETTEL VHENRNGTYQMIYKPSVNYQFKNTITVP 60			
DB	1 MFENPFKESPNVKTETETIQSIVVNYETTEL VHENRNGTYQMIYKPSVNYETFKNTIHP 60			
QY	61 KLGVMLVGMGNGNSSTLTGVIANREDI SWATKXKIQQANTYFGSLTQASAIRVSGFCEE 120			
DB	61 KLGVMLVGMGNGNSSTLTGVIANREGISWATKXKIQOANYFGSLTQASAIRVSGFCEE 120			
QY	121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPVMSWVPLGI 180			
DB	121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPVMSWVPLGI 180			
QY	181 YDPPFIANQBERANNVIGTKQEOVOQIIKIDIRAFKATKVDKVVVLMNTANTERSNLV 240			
DB	181 YDPPFIANQBERANNVIGTKQEOVOQIIKIDIRAFKATKVDKVVVLMNTANTERSNLV 240			
QY	241 VGLNDYMENTLAAVDRNEAETSPSTLYAIACVMENVPFINSPONTFVPGIIDLAIARNT 300			
DB	241 VGLNDYMENTLAAVDRNEAETSPSTLYAIACVMENVPFINSPONTFVPGIIDLAIARNT 300			
QY	301 LIGDDPFSGQTKKSVLVDFLVAGIKPTSIIVSYNHLGNNIDGNLSAPOTFRSKETSKS 360			
DB	301 LIGDDPFSGQTKKSVLVDFLVAGIKPTSIIVSYNHLGNNIDGNLSAPOTFRSKETSKS 360			

```

Db 301 LIGGDDFKSGOTKMKSVLDFVGAGIKPTSIIVSNHLGNNDGNMLSAPOFRSEKISKS 360
Qy 361 NVVDMDVNSNAIIVEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEIFMGKNITVLHNTC 420
Db 361 NVVDMDVNSNAIIVEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEIFMGKNITVLHNTC 420
Qy 421 EDLSLAAPITIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTKAPLPVPGTPEVN 480
Db 421 EDLSLAAPITIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTKAPLPVPGTPEVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510

```

RESULT 2

```

Q94KU3 PRELIMINARY; PRT; 510 AA.
ID Q94KU3
AC Q94KU3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Myo-inositol-3-phosphate synthase.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21196082; PubMed=11299373; DOI=10.1104/PP.125.4.1941;
RA Hegeman C.E., Good L.L., Grabau E.A.;
RT "Expression of D-myo-inositol-3-phosphate synthase in soybean.
RT Implications for phytic acid biosynthesis.";
RL Plant Physiol. 125:1941-1948 (2001).
DR EMBL; AF293970; AAK49896.1; -.
DR HSSP; P11986; 1PIH.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
SQ SEQUENCE 510 AA; 56506 MW; DE4F3DD7DC6F370 CRC64;

```

```

Query Match 98.1%; Score 2580; DB 2; Length 510;
Best Local Similarity 98.0%; Pred. No. 1.9e-161;
Matches 500; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 1 MFIEIFKESPNVKTETETISQVNVYETTELVEHNRNGTYQWIVKPKSVNYQFKTNTHP 60
Db 1 MFIEIFKESPNVKTETETISQVNVYETTELVEHNRNGTYQWIVKPKSVNYQFKTNTHP 60
Qy 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVSGFOGEE 120
Db 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVSGFOGEE 120
Qy 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMAAKAFVDIDLQOLRPYEWSPVLPFGI 180
Db 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMAAKAFVDIDLQOLRPYEWSPVLPFGI 180
Qy 181 YDPPFIANOBERANVVKGTQKQOVQOIIKDIKAFKATKVDKVVVMTANTERYSNLV 240
Db 181 YDPPFIANOBERANVVKGTQKQOVQOIIKDIKAFKATKVDKVVVMTANTERYSNLV 240
Qy 241 VGLNDTMENLLAADRNEAISPSTLYAIACWENVPFINSPOPTFVGLIDLAIANT 300
Db 241 VGLNDTMENLLAADRNEAISPSTLYAIACWENVPFINSPOPTFVGLIDLAIANT 300
Qy 301 LIGGDDFKSGOTKMKSVLDFVGAGIKPTSIIVSNHLGNNDGNMLSAPOFRSEKISKS 360
Db 301 LIGGDDFKSGOTKMKSVLDFVGAGIKPTSIIVSNHLGNNDGNMLSAPOFRSEKISKS 360

```

```

Qy 361 NVVDMDVNSNAIIVEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEIFMGKNITVLHNTC 420
Db 361 NVVDMDVNSNAIIVEPGEHPDHVVVIKYVPYVGDSSKRAMDEYTSIEIFMGKNITVLHNTC 420
Qy 421 EDLSLAAPITIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTKAPLPVPGTPEVN 480
Db 421 EDLSLAAPITIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTKAPLPVPGTPEVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510

```

RESULT 3

```

INOL_TOBAC STANDARD; PRT; 510 AA.
ID INOL_TOBAC
AC Q9LW96;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
DE synthase) (M1-1-P synthase) (IPS).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=leaf;
RX MEDLINE=20399434; PubMed=10945337;
RA Hara K., Yagi M., Koizumi N., Kusano T., Sano H.;
RT "Screening of wound-responsive genes identifies an immediate-early
RT expressed gene encoding a highly charged protein in mechanically
RT wounded tobacco plants.";
RL Plant Cell Physiol. 41:684-691 (2000).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = ID-myo-inositol 3-
CC phosphate.
CC -1- COPACTOR: NAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.

```

```

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

```

```

DR EMBL; AB009881; BAA95788.1; -.
DR HSSP; P11986; 1PIH.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56369 MW; 4BA8FDDA5DBF6DAD CRC64;

```

```

Query Match 94.1%; Score 2475; DB 1; Length 510;
Best Local Similarity 92.7%; Pred. No. 1.5e-154;
Matches 473; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

```

```

Qy 1 MFIEIFKESPNVKTETETISQVNVYETTELVEHNRNGTYQWIVKPKSVNYQFKTNTHP 60
Db 1 MFIEIFKESPNVKTETETISQVNVYETTELVEHNRNGTYQWIVKPKSVNYQFKTNTHP 60
Qy 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVSGFOGEE 120
Db 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVSGFOGEE 120
Qy 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMAAKAFVDIDLQOLRPYEWSPVLPFGI 180
Db 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMAAKAFVDIDLQOLRPYEWSPVLPFGI 180

```



```

Db 121 IYAPFKSLPMVNPDDVVFSGWMDISGMNLADAMARAKVFDIDLQKQLRPYMESWVPLPGI 180
Qy 181 YPDDTIANQGERANNVKGTQKQOVQOIKDKIKAFKATKVDKVVVLTANTERSNV 240
Db 181 YPDDTIANQGERANNVKGTQKQOVQOIKDKIKAFKATKVDKVVVLTANTERSNV 240
Qy 241 VGLNTMENTLAAVDRNEAISPSTLYAIACVMEVVPFNGSPONTFVGLIDLAIARNT 300
Db 241 VGLNTMENTLAAVDRNEAISPSTLYAIACVMEVVPFNGSPONTFVGLIDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMSVLDVFLVAGIKRPTSVSYNHLGNDGMNLSAPQTRSKSISKS 360
Db 301 LIGGDDFKSGQTKMSVLDVFLVAGIKRPTSVSYNHLGNDGMNLSAPQTRSKSISKS 360
Qy 361 NVYDDMANSNALYERGEHPDHVVVVKYVPYVGDSCRAMDETSBIFMGKNTIVLHNTC 420
Db 361 NVYDDMANSNALYERGEHPDHVVVVKYVPYVGDSCRAMDETSBIFMGKNTIVLHNTC 420
Qy 421 EDSLAAPIIIDVLVLAELSTRIOFKAENEGKFSFHVATILSYLTAKPLVPPGTPVNV 480
Db 421 EDSLAAPIIIDVLVLAELSTRIOFKAENEGKFSFHVATILSYLTAKPLVPPGTPVNV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

RESULT 4

```

IN01_NICPA STANDARD; PRT; 510 AA.
AC 09SV4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
synthase) (MI-1-P synthase) (IPS).
GN Name=INPS1;
OS Nicotiana paniculata.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=62141;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto A., Yamada S., Komori T.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
phosphate.
CC -1- CORFACTOR: NAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB032073; BAA04084.1; -.
CC DR HSSP; P11986; 1PIH.
CC DR InterPro; IPR002587; Inos-1-P synth.
CC DR Pfam; PF01658; Inos-1-P synth. 1.
CC KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
CC SEQUENCE 510 Aa; 56385 Mw; 41581c27a267666 CRC64;

```

```

Query Match 93.9%; Score 2470; DB 1; Length 510;
Best Local Similarity 92.4%; Pred. No. 3, 2e-154;
Matches 471; Conservative 21; Mismatches 16; Indels 0; Gaps 0;

```

```

Qy 1 MFIENPKVESPVVKYETELETIOSVYNETTELHENRNGTYOWI VPKSVNYQKNTIHPV 60
Db 1 MFIENPKVESPVVKYETELETIOSVYNETTELHENRNGTYOWI VPKSVNYQKNTIHPV 60
Qy 61 KLGWNLVGMGNGNGSTLGGVIANREDISWATKDKIQOANVGSYLTOASAIRVSGFQGBE 120
Db 61 KLGWNLVGMGNGNGSTLGGVIANREDISWATKDKIQOANVGSYLTOASAIRVSGFQGBE 120
Qy 121 IYAPFKSLPMVNPDDVVFSGWMDISGMNLADAMARAKVFDIDLQKQLRPYMESWVPLPGI 180
Db 121 IYAPFKSLPMVNPDDVVFSGWMDISGMNLADAMARAKVFDIDLQKQLRPYMESWVPLPGI 180
Qy 181 YPDDTIANQGERANNVKGTQKQOVQOIKDKIKAFKATKVDKVVVLTANTERSNV 240
Db 181 YPDDTIANQGERANNVKGTQKQOVQOIKDKIKAFKATKVDKVVVLTANTERSNV 240
Qy 241 VGLNTMENTLAAVDRNEAISPSTLYAIACVMEVVPFNGSPONTFVGLIDLAIARNT 300
Db 241 VGLNTMENTLAAVDRNEAISPSTLYAIACVMEVVPFNGSPONTFVGLIDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMSVLDVFLVAGIKRPTSVSYNHLGNDGMNLSAPQTRSKSISKS 360
Db 301 LIGGDDFKSGQTKMSVLDVFLVAGIKRPTSVSYNHLGNDGMNLSAPQTRSKSISKS 360
Qy 361 NVYDDMANSNALYERGEHPDHVVVVKYVPYVGDSCRAMDETSBIFMGKNTIVLHNTC 420
Db 361 NVYDDMANSNALYERGEHPDHVVVVKYVPYVGDSCRAMDETSBIFMGKNTIVLHNTC 420
Qy 421 EDSLAAPIIIDVLVLAELSTRIOFKAENEGKFSFHVATILSYLTAKPLVPPGTPVNV 480
Db 421 EDSLAAPIIIDVLVLAELSTRIOFKAENEGKFSFHVATILSYLTAKPLVPPGTPVNV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

RESULT 5

```

IN01_SESIN STANDARD; PRT; 510 AA.
AC 09FV1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
synthase) (MI-1-P synthase) (IPS).
OS Sesamum indicum (Oriental sesame) (Gingelly).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiales; Pedaliaceae; Sesamum.
OX NCBI_TaxID=4182;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Seed;
RC jin U.-H., Chung C.-H.;
RT "Characterization and functional analysis of a myo-inositol 1-
phosphate synthase cDNA from sesame (Sesamum indicum L.) seeds.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
phosphate.
CC -1- CORFACTOR: NAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```
CC -----
DR EMBL; AF284065; AAC0148.1; -.
DR HSSP; P11986; 1PIK.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56234 MW; 88D75376CE73401F CRC64;

Query Match 93.2%; Score 2452; DB 1; Length 510;
Best Local Similarity 92.2%; Pred. No. 5e-153;
Matches 470; Conservative 20; Mismatches 20; Indels 0; Gaps 0;

QY 1 MFENFKVESPNVXYTETETIOSVYNYETTELVHNRNGTYOMIVKPKSVNYQFKTNTHVP 60
DB 1 MFESFKVESPNVXYTETETIOSVYNYETTELVHNRNGTYOMIVKPKSVNYQFKTNTHVP 60
QY 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFSGLTQASAIRVGSFQEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFSGLTQASAIRVGSFQEE 120
QY 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
DB 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
QY 181 YDPPFIANQBERANNVIKGTKEQVOQIIKDIAKFEATKVDKVVLMTANTERYSNLV 240
DB 181 YDPPFIANQBERANNVIKGTKEQVOQIIKDIAKFEATKVDKVVLMTANTERYSNLV 240
QY 241 VGLNDTMENLLAADRNEAETSPSTLYAIACVMEVNPFGINSPONTFVPGILDLAIANT 300
DB 241 VGLNDTMENLLAADRNEAETSPSTLYAIACVMEVNPFGINSPONTFVPGILDLAIANT 300
QY 301 LIIGDDFFSGQTKMKSVLVDFVAGAGIKPTSIIVYNHLGNDGNMLSAPQTFRSKEISKS 360
DB 301 LIIGDDFFSGQTKMKSVLVDFVAGAGIKPTSIIVYNHLGNDGNMLSAPQTFRSKEISKS 360
QY 361 NVDDDMVNSAILYEPGHPDHVVVVKYVPYVGSKRAMEYTSSEIFMGKNITVLAHNTC 420
DB 361 NVDDDMVNSAILYEPGHPDHVVVVKYVPYVGSKRAMEYTSSEIFMGKNITVLAHNTC 420
QY 421 EDLSLAAPITIIDVLLAELSTRIOFKAENEGKFSFHPVATITLSYLTQAPLVPGTTPVN 480
DB 421 EDLSLAAPITIIDVLLAELSTRIOFKAENEGKFSFHPVATITLSYLTQAPLVPGTTPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 6
INOL WHEAT STANDARD; PRT; 510 AA.
ID _INOL WHEAT
AC Q9STU0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (Mt-1-P synthase) (IPS).
GN Name=MIPS;
OS Triticum aestivum (wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
OC Triticaceae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Fielder, and cv. Taber;
RA Hussain A., Yan W., Bock C., Baga M., Chibbar R., Georges F.;
RT "cDNA clone for myo-inositol 1-phosphate synthase from wheat.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = ID-myo-inositol 3-phosphate.
CC -1- COFACTOR: NAD (By similarity).
```

```
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements or send an email to license@sb-sib.ch).
CC -----
DR EMBL; AF120146; AAD26330.1; -.
DR EMBL; AF120147; AAD26331.1; -.
DR EMBL; AF120148; AAD26332.1; -.
DR HSSP; P11986; 1PIJ.
DR InterPro; IPR002587; Inos-1-P_synth.
DR Pfam; PF01658; Inos-1-P_synth; 1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56294 MW; 497D959C4456C37 CRC64;

Query Match 91.2%; Score 2400; DB 1; Length 510;
Best Local Similarity 89.4%; Pred. No. 1.3e-149;
Matches 456; Conservative 32; Mismatches 22; Indels 0; Gaps 0;

QY 1 MFENFKVESPNVXYTETETIOSVYNYETTELVHNRNGTYOMIVKPKSVNYQFKTNTHVP 60
DB 1 MFESFKVESPNVXYTETETIOSVYNYETTELVHNRNGTYOMIVKPKSVNYQFKTNTHVP 60
QY 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFSGLTQASAIRVGSFQEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFSGLTQASAIRVGSFQEE 120
QY 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
DB 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
QY 181 YDPPFIANQBERANNVIKGTKEQVOQIIKDIAKFEATKVDKVVLMTANTERYSNLV 240
DB 181 YDPPFIANQBERANNVIKGTKEQVOQIIKDIAKFEATKVDKVVLMTANTERYSNLV 240
QY 241 VGLNDTMENLLAADRNEAETSPSTLYAIACVMEVNPFGINSPONTFVPGILDLAIANT 300
DB 241 VGLNDTMENLLAADRNEAETSPSTLYAIACVMEVNPFGINSPONTFVPGILDLAIANT 300
QY 301 LIIGDDFFSGQTKMKSVLVDFVAGAGIKPTSIIVYNHLGNDGNMLSAPQTFRSKEISKS 360
DB 301 LIIGDDFFSGQTKMKSVLVDFVAGAGIKPTSIIVYNHLGNDGNMLSAPQTFRSKEISKS 360
QY 361 NVDDDMVNSAILYEPGHPDHVVVVKYVPYVGSKRAMEYTSSEIFMGKNITVLAHNTC 420
DB 361 NVDDDMVNSAILYEPGHPDHVVVVKYVPYVGSKRAMEYTSSEIFMGKNITVLAHNTC 420
QY 421 EDLSLAAPITIIDVLLAELSTRIOFKAENEGKFSFHPVATITLSYLTQAPLVPGTTPVN 480
DB 421 EDLSLAAPITIIDVLLAELSTRIOFKAENEGKFSFHPVATITLSYLTQAPLVPGTTPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 7
Q944C3 PRELIMINARY; PRT; 510 AA.
ID Q944C3;
AC Q944C3;
DT 01-DEC-2001 (TRENBLUREl. 19, Created)
DT 01-DEC-2001 (TRENBLUREl. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLUREl. 24, Last annotation update)
DE Myo-inositol-1-phosphate synthase.
GN Name=MIPS;
OS Suaeda salsa.
```

Query Match	91.2%	Score 2399	DB 2	Length 510
Best Local Similarity	90.4%	Pred. No. 1.5e-149		
Matches 462	Conservative 22	Mismatches 15	Indels 2	Gaps 2
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Amaranthaceae; Suaeda.				
OX NCBI_TaxID=126914				
RN [1]				
RP SEQUENCE FROM N.A.				
RA Wang L.P., Wang F.P., Sun Y.F., Zhao Y.X., Zhang H.;				
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.				
DR EMBL; AF433879; AAL28131.1; -				
DR HSBP; p1986; 1p1K.				
DR GO; GO:0004512; P:inositol-3-phosphate synthase activity; IEA.				
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.				
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.				
DR InterPro; IPR002587; Inos-1-P_synth.				
DR Pfam; PF01658; Inos-1-P_synth; 1.				
SQ SEQUENCE 510 AA; 56665 MW; 6C31006D2E1C508A CRC64;				
QY 1 MFIEHFKESPVVKVKTETETIOSVNVYETTELVENRNGV-YQWIKPKSVNVQPTNTNV				59
DB 1 MFIEHFKESPVNVYTNENSHSVNYETTELVENRKSQGYQWIKPKVTKYQFTQTQHV				60
QY 60 PGLGVMLVGMGGNNSTLTGVIANREDISWATKDKIQDANYFGLTQASAIRVSGFOGE				119
DB 61 PGLGVMLVGMGGNNCTLTGVIANREGISWATKDKIQDANYFGLTQASAIRVSGFNGE				120
QY 120 EIVAFPKSLPVMVNDIVYFGGWDISNNMLADMAPAKVFDIDLOKLRPMESVPLPG				179
DB 121 EIVAFPKSLPVMVNDIVYFGGWDISDLNADAMTRAKVLDIDOKLRPYMEHVPPLPG				180
QY 180 IYDPPEFIANDOEERANNVYIKGTQKOEVOQIIKDIAFKPATVVDKVVVLTANTERYSNL				239
DB 181 IYDPPEFIANDQSRANNIYIKGTKEQVEQVLDIIEFEKKNKVDVVVLTANTERYSVDV				240
QY 240 VVGLNDTMENTLAAVDRNEABISPESTLYAICWENVEPFINGSPONTFVPGILDIADARN				299
DB 241 VVGLNDTMENTLAAVDRNEABISPESTLYAICIGDVEPFINGSPONTFVPGILDIADARN				300
QY 300 TLIGGDDPKSGGTTKKKSVLVDFLVGAGIKFPISTIVSYNHLGNNDGNLSAPQTFRSKEISK				359
DB 301 CLIGGDDPKSGGTTKKKSVLVDFLVGAGIKFPISTIVSYNHLGNNDGNLSAPQTFRSKEISK				360
QY 360 SNVVDMMNSNALLYEPHEPHVNVVYKVPVVGSKRAMDEYSEIFMGGKNTVILANT				419
DB 361 SNVVDMMNSNALLYEPHEPHVNVVYKVPVVGSKRAMDEYSEIFMGGKNTVILANT				420
QY 420 CEDSLAAPIIIDLVLAELSTRIOFKAKENEGKFSFHPVATILSYLTKAPLVPGTTPV				479
DB 421 CEDSLAAPIIIDLVLAELSTRIOFKAKETBEKFTTFHPVATILSYLTKAPLVPGTTPV				480
QY 480 NALSKORAMLENIMRACVGLAENNMILEYK 510				
DB 481 NALSKORAMLEK-MRACVGLAENNMILEYK 510				
RESULT 8				
INOI_MESCR STANDARD; PRT; 512 AA.				
AC 040271;				
DT 16-OCT-2001 (Rel. 40, Last sequence update)				
DT 16-OCT-2001 (Rel. 40, Last sequence update)				
DT 28-FEB-2003 (Rel. 41, Last annotation update)				
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate synthase) (Mr-1-P synthase) (1PS).				
OS Mesembryanthemum crataegillum (Common ice plant).				
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
OC Caryophyllales; Aizoaceae; Mesembryanthemum.				
OX NCBI_TaxID=3544;				
RN [1]				

```

RP SEQUENCE FROM N.A.
RX MEDLINE=96209959; PubMed=8624516;
RA Ishlanti M., Majumder A.L., Bornhouser A., Michalowski C.B.,
RA Jensen R.G., Bohner H.J.;
RT "Coordinate transcriptional induction of myo-inositol metabolism
RT during environmental stress.";
RL Plant J. 9:537-548 (1996).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
CC phosphate.
CC -1- COFACTOR: NAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcement/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; U32511; AAB03687.1; -.
DR PIR; T12438; T12438.
DR HSSP; P11986; P1K.
DR InterPro; IPR002587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth. 1.
DR Inositol biosynthesis; Isomerase; NAD, Phospholipid biosynthesis.
SQ SEQUENCE 512 AA; 56758 MW; D4AB109FF47E8516 CRC64;

```

Query Match	91.0%;	Score 2395;	DB 1;	Length 513;
Best Local Similarity	89.5%;	Pred. No.295-149;		
Matches	458;	Conservative	27;	Mismatches 25; Indels 2; Gaps 1;
QY	1	MFIEFKYESNNVKTETETIOSVYNNYTTBLVHNNR--GTGYMIVKPKSVNTQFKNTNTH	58	
DB	1	MFISFKVESPNVKTETENIESVYNNYTTTELVHENRDKAGGYQWIVKPKVQVHFKTDTR	60	
QY	59	VPKIGWMLVGGNGNSTLGGVIANNEEDISMAFKDKIOQANYFGSLTOASAIRVGSFOG	118	
DB	61	VPKIGWMLVGGNGNSTLGGVIANNEEGISMAFKDKIOQANYFGSLTOASSIRVGSFNG	120	
QY	119	EEIYAPFKSLLPMVNPDDIVFGGNDISNNMLADAMARAKVFIDLOKOLAPYMSNVLP	178	
DB	121	EEIYAPFKSLLPMVNPDDIVFGGNDISDNMLADAMTARVFIDLOKOLAPYMEHWPPLP	180	
QY	179	GIYPPDPFIANQEBRANNVIKGTQEOBOVOOIIDIKAFKATVYDKVYLMTANTREYSN	238	
DB	181	GIYPPDPFIANQGBRANNVIKGTKEVEVERITKDIRFEKKKNKVDKVVVLMTGTEYSN	240	
QY	239	LVVGLGNTDMENLLAAVDRNEAETISPTLYAIACYMEVNPFIINGSPONTFVPGILDILAAR	298	
DB	241	VVGLGNTDMENLLASLEKNESEISPSISYALACTEENIPIINGSPONTFVPGILDILAARK	300	
QY	299	NTLIGGDDFKSGGQTKMKSVLVDPLVGAIGIKPTSIIVSYNHLGNDGNMLSAPQTFRSKEIS	358	
DB	301	NSLIGGDDFKSGGQTKMKSVLVDPLVGAIGIKPTSIIVSYNHLGNDGNMLSAPQTFRSKEIS	360	
QY	359	KSNVYDDMVNSNALIYEPGEHPHVAVVYIKVVPYGDGSKRAMDEYTSIEIPAGGKNTIIVLHN	418	
DB	361	KSNVYDDMVNSNGLIYEPGEHPHVAVVYIKVVPYGDGSKRAMDEYTSIEIPAGGKNTIIVLHN	420	
QY	419	TCEDSLLAAPITLDVLLAELSTRIOFKAENEGKFNHFHVPATILSYLTAAPLVPPGTPV	478	
DB	421	TCEDSLLAAPITLDVLLAELSTRIOFKAEDEBKFNHFHVPATILSYLTAAPLVPPGTPV	480	
QY	479	VNALSKORAMENIMRACVGLAPENNMTILEYK	510	
DB	481	VNALSKORAMENILIRACVGLAPENNMTILEYK	512	

```

IN02 ARATH STANDARD; PRT; 510 AA.
ID IN02 ARATH
AC Q38862; Q38862;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Inositol-3-phosphate synthase isozyme 2 (EC 5.5.1.4) (Myo-inositol-1-
phosphate synthase 2) (MI-1-P synthase 2) (IPS 2).
GN OrderedLocNames=At2g22240;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RN RP
RP SEQUENCE FROM N.A.
RC STRAIN=cnv. Columbia;
RA Johnson M.D., Burk D.H.;
RT "Isozyme of 1L-myo-inositol-1-phosphate synthase from Arabidopsis.",
RL (ex) Plant Gene Register PGR95-067.
[2]
RN RP
RP SEQUENCE FROM N.A.
RC STRAIN=cnv. Columbia;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Rongning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umeyam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Cressy T.H.,
RA Goodman H.M., Somerville C.R., Copenhagen G.P., Preuss D.,
RA Niemman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768(1999).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
phosphate.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U30250; AAC49172.1; -.
DR EMBL; AC007168; AAD23618.1; -.
DR PIR; D84610; D84610.
DR HSSP; P11986; P119.
DR InterPro; IPR002587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth; 1.
KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
FT CONFLICT 135 135 E -> D (in Ref. 1).
FT CONFLICT 287 287 K -> N (in Ref. 1).
FT CONFLICT 298 298 K -> N (in Ref. 1).
FT CONFLICT 318 318 L -> W (in Ref. 1).
FT CONFLICT 477 477 P -> A (in Ref. 1).
FT CONFLICT 487 487 A -> P (in Ref. 1).
SQ SEQUENCE 510 AA; 56337 MW; BDD58C1668289CES CRC64;

Query Match 90.6%; Score 2384; DB 1; Length 510;
Best Local Similarity 88.2%; Pred. No. 1,5e-148;
Matches 450; Conservative 37; Mismatches 23; Indels 0; Gaps 0;

QY 1 MFIEFKVESPVKKTETETISGVNYETTELTHENRNGYQWIVKPSVNYQKTNTHVP 60
DB 1 MFIEFVKVESPVKKTETENINSYDYTEVTEVHENRNGYQWVVPKTKVDFKDTITVP 60

```

```

QY 61 KLGVMVLVGMGNGNSTLTGVIANREDI SWATKDKIQANYFGSLTQASAIRVSGFOCEE 120
DB 61 KLGVMVLVGMGNGNSTLTGVIANKEGISWATKDKIQOANYFGSLTQASSIRVSGSYNEE 120
QY 121 IYAFPKSLTPVNVNDDIVFGGWDISNNMLADAMARAVFDIDLOKOLRPVNESVNPJGI 180
DB 121 IYAFPKSLTPVNVNDDIVFGGWDISNNMLADAMARAVFDIDLOKOLRPVNEENIPLPFI 180
QY 181 YDPPFIANQGERANNVYKGTQEOVOOI IKDIAFKFATKVDKVVVLMNTANRYSNLV 240
DB 181 YDPPFIANQGSRRANSYIKGTKEQVDHIIKDMEFKKNVDLVVMTANRYSNVI 240
QY 241 VGLNDTMENTLLAAVDRNEAEIISPSTLYAIACWMENVPPINGS PONTFVPGIIDLAIARNT 300
DB 241 VGLNDTMENTLLAAVDEKDESEISPSTLYAIACVLGIPFINGS PONTFVPGIIEAISNKC 300
QY 301 LIIGDDPKSGQTKKSVYVDFLVGAGIKPRTSVSYNNHGNNDGNLSPQFRSKETSKS 360
DB 301 LIIGDDPKSGQTKKSVYVDFLVGAGIKPRTSVSYNNHGNNDGNLSPQFRSKETSKS 360
QY 361 NVVDMMVNSNAIIXEPGEHPDHVVYIKVVPYVDSKRAMDEYTSIEIFMGKNTITVLNHTC 420
DB 361 NVVDMMVNSNAIIXEPGEHPDHVVYIKVVPYVDSKRAMDEYTSIEIFMGKNTITVLNHTC 420
QY 421 EDSLIAPITIDVLVLAELSTRIOFKANEGKFSFHPVATILSYTKAPLVPGTPVYN 480
DB 421 EDSLIAPITIDVLVLAELSTRIOFKANEGKFSFHPVATILSYTKAPLVPGTPVYN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMIIEYK 510
DB 481 ALSKORAMLENILRACVGLAPENNMIIEYK 510

RESULT 10
IN01 BRANA STANDARD; PRT; 510 AA.
ID IN01 BRANA
AC Q96348;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
synthase) (MI-1-P synthase) (IPS).
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
[1]
RN RP
RP SEQUENCE FROM N.A.
RA Huesain A., Bourgeois J., Poly S., Tsang E., Keller W.A., Georges F.;
RT "Cloning of a full length cDNA encoding myo-inositol 1-phosphate
synthase from Brassica napus ";
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
phosphate.
CC -1- COFACTOR: NAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U66307; AAB06756.2; -.
DR PIR; T08436; T08436.
DR HSSP; P11986; P119.

```

DR InterPro; IPR002587; Inos-1-P synth.
 KW Pfam; PF01658; Inos-1-P synth. 1.
 DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
 SQ SEQUENCE 510 AA; 56377 MW; A40EB6558D80739 CRC64;

Query Match 90.3%; Score 2376; DB 1; Length 510;
 Best Local Similarity 88.6%; Pred. No. 5e-148;
 Matches 452; Conservative 29; Mismatches 29; Indels 0; Gaps 0;

```

QY 1 MFIEPFKESPNVKTETETIEIOSVNYETTELVEHNRNGTYQVIYKPKSVNTQFKTNHVP 60
DB 1 MFIEPFKESPNVKTETETIEIOSVNYETTELVEHNRNGTYQVIYKPKSVNTQFKTNHVP 60
QY 61 KLGVNLVGMGNGNSTLTAGVIANREGISWATKDKVQOANFGSLTQASIRVGSFQEE 120
DB 61 KLGVNLVGMGNGNSTLTAGVIANREGISWATKDKVQOANFGSLTQASIRVGSFQEE 120
QY 121 IYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKOLRPYMESVPLPGI 180
DB 121 IYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKOLRPYMESVPLPGI 180
QY 121 MYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKOLRPYMESVPLPGI 180
DB 121 MYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKOLRPYMESVPLPGI 180
QY 181 YDPDFIAANQGBRANNVIKGTKEQVOQIIKDIIKAFKATKVDKVVVMTANTERYSNLV 240
DB 181 YDPDFIAANQGBRANNVIKGTKEQVOQIIKDIIKAFKATKVDKVVVMTANTERYSNLV 240
QY 241 VGLNDTMENLLAANDRAEISPSLTVAIACVMEVNPFIKSPQTPVPGILDLAIARNT 300
DB 241 VGLNDTMENLLAANDRAEISPSLTVAIACVMEVNPFIKSPQTPVPGILDLAIARNT 300
QY 301 LIGGDDPFSGGOTKMSVLDVFLVAGIKPSTISVYNHLGNNDGNLSPQTFRSKEISKS 360
DB 301 LIGGDDPFSGGOTKMSVLDVFLVAGIKPSTISVYNHLGNNDGNLSPQTFRSKEISKS 360
QY 361 NVVDPMVNSNAILVEPGEHPHVIVIKYVPYVDSKRAMDEYTSIEMGKNTIVLHNTC 420
DB 361 NVVDPMVNSNAILVEPGEHPHVIVIKYVPYVDSKRAMDEYTSIEMGKNTIVLHNTC 420
QY 421 EDLSLLAAPIIIDVLVLAESTRIQKANEKGKFSFHVATILSLTQAPLVPRCTPVVN 480
DB 421 EDLSLLAAPIIIDVLVLAESTRIQKANEKGKFSFHVATILSLTQAPLVPRCTPVVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

```

RESULT 11

INOS_1 ARATH STANDARD; PRT; 510 AA.

```

AC 09LX12.16-0CT-2001 (Rel. 40, Created)
DT 16-0CT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Probable inositol-3-phosphate synthase isoform 3 (BC 5.5.1.4) (Myo-
DS inositol-1-phosphate synthase 3) (MT-1-P synthase 3) (IFS 3).
CN OrderedLocustNames=AL5G10170; ORFNames=T31P16_160;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016721; PubMed=1130714; DOI=10.1038/35048507;
RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu B.,
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA Kohara M., Katsumoto K., Matsuno A., Muraki A., Nakayama S.,
RA Nakazaki N., Naito K., Okumura S., Shitipo S., Takeuchi C., Wada T.,
RA Wakemabe A., Yamada M., Sato S., de la Bastide M.,
RA Huang B., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Stonestring T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,

```

```

RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
RA Leonard S., Meyer R., Molvaney E., Ozersky P., Riley A., Strommatt C.,
RA Wagner-McPherson C., Wolian A., Yoakum M., Bell W., Dedha N.,
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vail D., Baker J.,
RA Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M.A.,
RA Marienssen R., McCombe W.R., Wilson R.K., Murphy G., Bancroft I.,
RA Volckaert G., Wandurt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA Eutlan K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
RA Langham S.-A., McCullagh B., Robben J., Grynoprez B., Zimmermann W.,
RA Ramberger U., Wedler H., Balke K., Wedler E., Peters S.,
RA van Steveren M., Dirkse W., Mooijman P., Klein lankhorst R.,
RA Welzenegger T., Bothe G., Rose M., Hauf J., Bernaisier S., Hempel S.,
RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rüd S., Schoof H.,
RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
RT thaliana."
RL Nature 408:823-826(2000).
CC -I- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
CC phosphate.
CC -I- COFACTOR: NAD (By similarity).
CC -I- PATHWAY: Inositol biosynthesis.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AL356332; CAB92058.1; -.
DR PIR; T50021; T50021.
DR HSSD; P11986; 1PIH.
DR InterPro; IPR002587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth. 1.
KW Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56417 MW; 5CB8108082152473 CRC64;

```

Query Match 90.3%; Score 2376; DB 1; Length 510;
 Best Local Similarity 87.8%; Pred. No. 5e-148;
 Matches 448; Conservative 40; Mismatches 22; Indels 0; Gaps 0;

```

QY 1 MFIEPFKESPNVKTETETIEIOSVNYETTELVEHNRNGTYQVIYKPKSVNTQFKTNHVP 60
DB 1 MFIEPFKESPNVKTETETIEIOSVNYETTELVEHNRNGTYQVIYKPKSVNTQFKTNHVP 60
QY 61 KLGVNLVGMGNGNSTLTAGVIANREGISWATKDKVQOANFGSLTQASIRVGSFQEE 120
DB 61 KLGVNLVGMGNGNSTLTAGVIANREGISWATKDKVQOANFGSLTQASIRVGSFQEE 120
QY 121 IYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKOLRPYMESVPLPGI 180
DB 121 IYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKOLRPYMESVPLPGI 180
QY 121 MYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKOLRPYMESVPLPGI 180
DB 121 MYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKOLRPYMESVPLPGI 180
QY 181 YDPDFIAANQGBRANNVIKGTKEQVOQIIKDIIKAFKATKVDKVVVMTANTERYSNLV 240
DB 181 YDPDFIAANQGBRANNVIKGTKEQVOQIIKDIIKAFKATKVDKVVVMTANTERYSNLV 240
QY 181 MYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKOLRPYMESVPLPGI 180
DB 181 MYAPFKSLPVMNPDDIVFGGWDISNMNLADAMARAKYFDIDLOKOLRPYMESVPLPGI 180
QY 241 VGLNDTMENLLAANDRAEISPSLTVAIACVMEVNPFIKSPQTPVPGILDLAIARNT 300
DB 241 VGLNDTMENLLAANDRAEISPSLTVAIACVMEVNPFIKSPQTPVPGILDLAIARNT 300
QY 301 LIGGDDPFSGGOTKMSVLDVFLVAGIKPSTISVYNHLGNNDGNLSPQTFRSKEISKS 360
DB 301 LIGGDDPFSGGOTKMSVLDVFLVAGIKPSTISVYNHLGNNDGNLSPQTFRSKEISKS 360
QY 361 NVVDPMVNSNAILVEPGEHPHVIVIKYVPYVDSKRAMDEYTSIEMGKNTIVLHNTC 420
DB 361 NVVDPMVNSNAILVEPGEHPHVIVIKYVPYVDSKRAMDEYTSIEMGKNTIVLHNTC 420

```

Dh 361 NVDDMVGNSGILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVMNTC 420
Qy 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Dh 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Dh 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510

RESULT 12

Q9AR12 PRELIMINARY; PRT; 509 AA.
ID Q9AR12
AC Q9AR12
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Myo-inositol 1-phosphate synthase.
OS Alicentia marina (Grey mangrove).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Acanthaceae; Acanthaceae; Inceritae sedis; Alicentia.
NCBI_TaxID=82927;
RN [1]
RP SEQUENCE FROM N.A.
RA Jithesh M.N., Parani M., Parida A.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY028259; AAK21969.1; -.
DR HSSP; P1986; 1PKI.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth_1.
SQ SEQUENCE 509 AA; 55978 MW; 23C8D354BAF3BD0F CRC64;

Query Match 90.2%; Score 2373.5; DB 2; Length 509;
Best Local Similarity 89.6%; Pred. No. 7.3e-148;
Matches 457; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

Qy 1 MFIEFKVESPNVKTETETELVHNRNGTQWIKPKSVNQFKNTHVP 60
Dh 1 MFIEFKVESPNVKTETETELVHNRNGTQWIKPKSVNQFKNTHVP 60
Dh 1 MFIEFKVESPNVKTETETELVHNRNGTQWIKPKSVNQFKNTHVP 60
Qy 1 KLGWLVGMGNGNSTLTGVIANREDISWATKDIOQANYGSLTQASAIRVGSFOGEE 120
Dh 1 KLGWLVGMGNGNSTLTGVIANREDISWATKDIOQANYGSLTQASAIRVGSFOGEE 120
Dh 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDIOQANYGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
Dh 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
Qy 181 YDPDFIAANQGEERANNVTKGTKEQVOQIITKDIAFKKATKVDKVVVLTANTERYSNLV 240
Dh 181 YDPDFIAANQGEERANNVTKGTKEQVOQIITKDIAFKKATKVDKVVVLTANTERYSNLV 240
Dh 181 YDPDFIAANQGEERANNVTKGTKEQVOQIITKDIAFKKATKVDKVVVLTANTERYSNLV 240
Qy 241 VGLNDTMENLLAAVDREAEISPSSTLYAIACVMEVNPFGINSPONTFVPGILDLAIANT 300
Dh 241 VGLNDTMENLLAAVDREAEISPSSTLYAIACVMEVNPFGINSPONTFVPGILDLAIANT 300
Dh 240 VGLNDTMENLLAAVDREAEISPSSTLYAIACVMEVNPFGINSPONTFVPGILDLAIANT 300
Qy 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGNLSAPQTFRSKEISKS 360
Dh 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGNLSAPQTFRSKEISKS 360
Dh 300 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGNLSAPQTFRSKEISKS 359
Qy 361 NVDDMVGNSGILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVMNTC 420
Dh 361 NVDDMVGNSGILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVMNTC 420
Dh 360 NVDDMVGNSGILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVMNTC 419
Qy 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Dh 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Dh 420 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 479

Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Dh 480 ALSKORAMLENIMRACVGLAPENNMLEYK 509

RESULT 13

Q7XUCO PRELIMINARY; PRT; 510 AA.
ID Q7XUCO
AC Q7XUCO;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Myo-inositol 1-phosphate synthase INOI.
OS Xerophyta viscosa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Velloziaceae; Xerophyta.
NCBI_TaxID=90708;
RN [1]
RP SEQUENCE FROM N.A.
RA Majee M., Majumder A.N.L., Munder S.G.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY323824; AAP85531.1; -.
DR HSSP; P1986; 1UKI.
DR GO; GO:0004512; F:inositol-3-phosphate synthase activity; IEA.
DR GO; GO:0006021; P:myo-inositol biosynthesis; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR002587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth_1.
SQ SEQUENCE 510 AA; 56234 MW; 5F92212651115A2A CRC64;

Query Match 89.7%; Score 2361; DB 2; Length 510;
Best Local Similarity 88.6%; Pred. No. 4.8e-147;
Matches 452; Conservative 28; Mismatches 30; Indels 0; Gaps 0;

Qy 1 MFIEFKVESPNVKTETETELVHNRNGTQWIKPKSVNQFKNTHVP 60
Dh 1 MFIEFKVESPNVKTETETELVHNRNGTQWIKPKSVNQFKNTHVP 60
Dh 1 MFIEFKVESPNVKTETETELVHNRNGTQWIKPKSVNQFKNTHVP 60
Qy 1 KLGWLVGMGNGNSTLTGVIANREDISWATKDIOQANYGSLTQASAIRVGSFOGEE 120
Dh 1 KLGWLVGMGNGNSTLTGVIANREDISWATKDIOQANYGSLTQASAIRVGSFOGEE 120
Dh 61 KLGWLVGMGNGNSTLTGVIANREDISWATKDIOQANYGSLTQASAIRVGSFOGEE 120
Qy 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
Dh 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
Dh 121 IYAPFKSLPMVNPDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESWVPLPGI 180
Qy 181 YDPDFIAANQGEERANNVTKGTKEQVOQIITKDIAFKKATKVDKVVVLTANTERYSNLV 240
Dh 181 YDPDFIAANQGEERANNVTKGTKEQVOQIITKDIAFKKATKVDKVVVLTANTERYSNLV 240
Dh 181 YDPDFIAANQGEERANNVTKGTKEQVOQIITKDIAFKKATKVDKVVVLTANTERYSNLV 240
Qy 241 VGLNDTMENLLAAVDREAEISPSSTLYAIACVMEVNPFGINSPONTFVPGILDLAIANT 300
Dh 241 VGLNDTMENLLAAVDREAEISPSSTLYAIACVMEVNPFGINSPONTFVPGILDLAIANT 300
Dh 241 VGLNDTMENLLAAVDREAEISPSSTLYAIACVMEVNPFGINSPONTFVPGILDLAIANT 300
Qy 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGNLSAPQTFRSKEISKS 360
Dh 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGNLSAPQTFRSKEISKS 360
Dh 301 LIGGDDFKSGOTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGNLSAPQTFRSKEISKS 360
Qy 361 NVDDMVGNSGILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVMNTC 420
Dh 361 NVDDMVGNSGILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVMNTC 420
Dh 361 NVDDMVGNSGILYEPGEHPDHVVVVKYPCVGSKRAMEYTSIEFMGGKNTIVMNTC 420
Qy 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Dh 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Dh 421 EDSLAAPIIIDVLVLAELSTRIOFKANEKGKFSFHPVATILSYLTAKAPLVPGTPEVNV 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510
Dh 481 ALSKORAMLENIMRACVGLAPENNMLEYK 510

RESULT 14

INOI_SPTPO

```

ID INOI SPIPO STANDARD; PRT; 510 AA.
AC P42803;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
DE synthase) (MI-1-P synthase) (IPS).
GN Name=Tur1;
OS Spirodela polyrrhiza (Giant duckweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Araceae; Lemnoidae;
OC Spirodela.
OX NCBI_TaxID=29656;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94035182; PubMed=8220483;
RA Smart C.C., Fleming A.J.;
RT "A plant gene with homology to D-myo-inositol-3-phosphate synthase is
RT rapidly and spatially up-regulated during an abscisic-acid-induced
RT morphogenic response in Spirodela polyrrhiza.";
RL Plant J. 4:279-293(1993).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
CC phosphate.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- INDUCTION: By abscisic acid (ABA).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z11693; CAAT7751.1; -
DR PIR; S60302; S60302.
DR HSSP; P11986; 1PIH.
DR InterPro; IPR002587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth; 1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
SQ SEQUENCE 510 AA; 56385 MW; 2D56D3666FC5E03C CRC64;

Query Match 89.6%; Score 2358; DB 1; Length 510;
Best Local Similarity 87.8%; Pred. No. 7,6e-147;
Matches 448; Conservative 37; Mismatches 25; Indels 0; Gaps 0;

QY 1 MFIEFKVESPNVKKYTEIEIQSVNYETTELVHENRNGTYOMIVKPKSVNYQFKTNTHVP 60
DB 1 MFIEKFRVSPVKKYGGDGEISVSYETTELVEHVRNGSYQWVVKPKSVYQFKTDTRVP 60
QY 61 KLGVNLVGMGGNGSTLTGVTANREDISWATKDIQGANYSGLTQASAPRVSGFOEE 120
DB 61 KLGVNLVGMGGNGSTLTGVTANREGISWVTKEXYQGANFVGSITQSSSIRVGSFNGBE 120
QY 121 IYAPFKSLLPVNPDPDIVFGGWDISNNMLADAMARAKYFDIDLOQLRPYMESWPLPGI 180
DB 121 IYAPFKSLLPVNPDPDIVFGGWDISNNMLADAMARAKYFDIDLOQLRPYMESWPLPGI 180
QY 121 IYAPFKSLLPVNPDPDIVFGGWDISNNMLADAMARAKYFDIDLOQLRPYMESWPLPGI 180
DB 121 IYAPFKSLLPVNPDPDIVFGGWDISNNMLADAMARAKYFDIDLOQLRPYMESWPLPGI 180
QY 181 YDPDPIAANOSERRANVITKGTQBOVOQIIKDIAFKENATKYDKVVLMTANTERYSNLV 240
DB 181 YDPDPIAANOSERRANVITKGTQBOVOQIIKDIAFKENATKYDKVVLMTANTERYSNLV 240
QY 181 YNPDPPIAANOSERRANVITKGTQBOVOQIIKDIAFKENATKYDKVVLMTANTERYSNLV 240
DB 181 YNPDPPIAANOSERRANVITKGTQBOVOQIIKDIAFKENATKYDKVVLMTANTERYSNLV 240
QY 241 VGLNDTMEMLAAVDRNEAETISPTLYAIACVMEVNPFIINGSPONTFVGLIDLIAANT 300
DB 241 VGLNDTMEMLAAVDRNEAETISPTLYAIACVMEVNPFIINGSPONTFVGLIDLIAANT 300
QY 241 VGLNDTMEMLAAVDRNEAETISPTLYAIACVMEVNPFIINGSPONTFVGLIDLIAANT 300
DB 241 VGLNDTMEMLAAVDRNEAETISPTLYAIACVMEVNPFIINGSPONTFVGLIDLIAANT 300
QY 301 LIGGDDFKSGGOTKMSVLVDPLVGAIGIKPTSIIVSYNHLGNDGNMLSAAPOTFRSKEISKS 360
DB 301 LIGGDDFKSGGOTKMSVLVDPLVGAIGIKPTSIIVSYNHLGNDGNMLSAAPOTFRSKEISKS 360

```

```

QY 361 NVVDDMANSNALTVEPGEHPHVIVIKYVPVYGDSCRAMDEYTSRIFMGKXTIVLHNTC 420
DB 361 NVVDDMANSNALTVEPGEHPHVIVIKYVPVYGDSCRAMDEYTSRIFMGKXTIVLHNTC 420
QY 421 EDSLAAPIIIDLVLLAEISTRIQPKANEKGKFSFHPVATILSYLTKAPLVPFGTPVYN 480
DB 421 EDSLAAPIIIDLVLLAEISTRIQPKANEKGKFSFHPVATILSYLTKAPLVPFGTPVYN 480
QY 481 ALSKORAMLENTIRACVGLAPENNMLEYK 510
DB 481 ALSKORAMLENTIRACVGLAPENNMLEYK 510

RESULT 15
ID INOI MAIZE STANDARD; PRT; 510 AA.
AC 09PFR7; 065196;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inositol-3-phosphate synthase (EC 5.5.1.4) (Myo-inositol-1-phosphate
DE synthase) (MI-1-P synthase) (IPS).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Cv. Early ACR; TISSUE=leaf;
RA Larson S.R., Raboy V.;
RT "Linkage mapping maize and barley myo-inositol 1-phosphate synthase
RT genes.";
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Shukla S., Vantsoi T.T.;
RT "Genomic sequence of maize myo-inositol 1-phosphate synthase gene.";
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate = 1D-myo-inositol 3-
CC phosphate.
CC -1- COFACTOR: NAD (By similarity).
CC -1- PATHWAY: Inositol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the myo-inositol-1-phosphate synthase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF056326; AAC15756.1; -
DR EMBL; AF323175; AAG40328.1; -
DR PIR; T01647; T01647.
DR HSSP; P11986; 1PIH.
DR InterPro; IPR002587; Inos-1-P synth.
DR Pfam; PF01658; Inos-1-P synth; 1.
DR Inositol biosynthesis; Isomerase; NAD; Phospholipid biosynthesis.
KW CONFLICT 18 M -> T (in Ref. 2).
FT CONFLICT 351 351 A -> T (in Ref. 2).
SQ SEQUENCE 510 AA; 56245 MW; DAB59BECF391CB6D CRC64;

Query Match 89.5%; Score 2354; DB 1; Length 510;
Best Local Similarity 88.2%; Pred. No. 1.4e-146;
Matches 450; Conservative 27; Mismatches 33; Indels 0; Gaps 0;

QY 1 MFIEFKVESPNVKKYTEIEIQSVNYETTELVHENRNGTYOMIVKPKSVNYQFKTNTHVP 60
DB 1 MFIEFKVESPNVKKYTEIEIQSVNYETTELVHENRNGTYOMIVKPKSVNYQFKTNTHVP 60

```

```

Db      1 MFESFRVSPHVRVGYPMIESEYRYDTTELVEHGKGASRWVVRPKSVKYNFRTRTAVP 60
Qy      61 KLGVLVGMGNGNSTLTGVTIANREDISWATKDIQOANYFGSLTQASAIRVGSFOGEE 120
Db      61 KLGVLVGMGNGNSTLTGVTIANREGISWATKDVQOANYGSLTQASTIRVGSYNSEE 120
Qy      121 IYAPFKSLIPVNPDDIVFGGWDISNMNLIADAMARAKVFDIDLQKLRPYMESWVPLPGI 180
Db      121 IYAPFKSLIPVNPDDIVFGGWDISNMNLIADSMTRAKVLIDLQKLRPYMESWVPLPGI 180
Qy      181 YDPPEIANOEERANNVVKGTQOEVOQI IKDIAFKKATKVDKVVLTANTERYSNLV 240
Db      181 YDPPEIAANQSRANSVIKGTKEQVEQI IKDIREFEKKNKVDKIVLMTANTERYSNVC 240
Qy      241 VGLNDTMENLAAVDRNEAEISPTLYAIAQVMEVVPFINGSPONTFVPGIIDLAIANT 300
Db      241 AGLNDTMENLAAVDKNEAEVSPSTLYAIAQVMEGVFPFINGSPONTFVPGIIDLAIKNNC 300
Qy      301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSI VSYNHLGNNDGMNLSAPQTFRSKEISKS 360
Db      301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSI VSYNHLGNNDGMNLSAPQAFRSKEISKS 360
Qy      361 NVUDDMVNSNAI LYERGEHPDHVVVVKYVPYVGD SKRAMDEYTSRI FMGKNTIVLHNTC 420
Db      361 NVUDDMVNSNAI LYERGEHPDHVVVVKYVPYVGD SKRAMDEYTSRI FMGKNTIVLHNTC 420
Qy      421 EDSLIAPIIIDLVLLAELSTRIOFKANEKGKHSFHPVATILSYLTAKAPLVPGTPIVYN 480
Db      421 EDSLIAPIIIDLVLLAELSTRIOFKANEKGKHSFHPVATILSYLTAKAPLVPGTPIVYN 480
Qy      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db      481 ALAKORAMLENIMRACVGLAPENNMILEYK 510

```

Search completed: June 7, 2005, 16:46:41
 Job time : 87 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	2395	91.0	512	2	T12138	inositol-3-phosphat
2	2384	90.6	510	2	D84610	probable myo-inos
3	2376	90.3	510	2	T50021	inositol-3-phosph
4	2371	90.1	509	2	T08436	inositol-3-phosph
5	2358	89.6	510	2	S60302	inositol-3-phosph
6	2354	89.5	510	2	T01647	inositol-3-phosph
7	2349.5	89.3	511	2	T05017	inositol-3-phosph
8	2338	88.9	511	2	T04399	inositol-3-phosph
9	2328.5	88.5	511	2	T10964	inositol-3-phosph
10	2304.5	87.6	507	2	S52448	inositol-3-phosph
11	1393.5	53.0	525	2	T18569	inositol-3-phosph
12	1389.5	52.8	520	2	S45452	inositol-3-phosph
13	1389.5	52.8	555	2	A30902	inositol-3-phosph
14	1312.5	49.9	430	2	T46317	hypothetical prote
15	291	11.1	388	2	T36191	probable myo-inos
16	284	10.8	392	2	A69474	myo-inositol-1-ph
17	233	8.9	417	2	T34830	probable secret
18	215	8.2	352	2	H72255	myo-inositol-1-ph
19	214	8.1	425	2	E71039	hypothetical prote
20	197	7.5	386	2	B75175	myo-inositol-1-ph
21	193	7.3	364	2	G70451	conserved hypoth
22	161	6.1	360	2	B90239	conserved hypoth
23	156	5.9	352	2	F72632	probable myo-inos
24	143.5	5.5	368	2	A69014	conserved hypoth
25	142	5.4	369	2	S72835	hypothetical prote
26	132	5.0	360	2	T36586	hypothetical prote
27	127.5	4.8	367	2	F70912	hypothetical prote
28	124.5	4.7	357	2	G95871	conserved hypoth
29	120.5	4.6	1993	2	AF1450	probable peptidog

30	118.5	4.5	853	2	G90555	trse-1-like protein
31	118	4.5	731	2	D95235	penicillin-binding
32	118	4.5	802	2	C90326	hypothetical prote
33	117.5	4.5	2481	2	A43908	fibronectin - Afri
34	116	4.4	1127	2	T28317	ORF MSV156 hypothe
35	115.5	4.4	452	2	S77155	UDP-N-acetylmuramo
36	115.5	4.4	873	2	T16282	hypothetical prote
37	114.5	4.4	344	2	B90176	conserved hypothet
38	114	4.3	731	2	E98099	peptidoglycan glyco
39	113.5	4.3	1628	2	E90538	hypothetical prote
40	113.5	4.3	1792	2	T20363	hypothetical prote
41	113	4.3	867	2	H90524	preprotein translo
42	112.5	4.3	900	2	S46101	hypothetical prote
43	112	4.3	2490	1	A54971	protein-tyrosine-p
44	111.5	4.2	422	2	C69953	isocitrate dehydro
45	111.5	4.2	867	2	B68815	ClpB protein (limp

ALIGNMENTS

RESULT 1

inosticol-3-phosphate synthase (EC 5.5.1.4) - common ice plant
C:Species: Mesembryanthemum crystallinum (common ice plant)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C:Accession: T12438
C:Right(s): M.; Majumder, A.L.; Bornhauser, A.; Michalowski, C.B.; Jensen, R.G.; Bohner
Plant J. 9, 537-548, 1996
A:Title: Coordinate transcriptional induction of myo-inosticol metabolism during environm
A:Reference number: Z17518, MID:96208959, PMID:8624516
A:Accession: T12438
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-512 <ISH>
A:Cross-references: UniProt:Q040271, EMBL:U32511, NID:g9755887, PIDD:AA03687.1, PID:g97558
C:Superfamily: myo-inosticol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase

Query Match 91.0%; Score 2395; DB 2; Length 512;

Matches 458; Conservative 27; Mismatches 25; Indels 2; Gaps 1.

[illegible]

```
Db 421 TCEDSLAAPIILDVLLAELSTRIOQLKAEBDEKFSFHPVATILSYLTAKALVPFGTPV 480
Qy 479 VNALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 VNALSKORAMLENIMRACVGLAPENNMILEYK 512

RESULT 2
Db4610
probable myo-inositol 1-phosphate synthase (imported) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Aug-2004
C:Accession: DB4610
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon, L.;
Eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-766, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A04420; MUID:20083487; PMID:10617197
A:Accession: DB4610
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-510 <STO>
A:Cross-references: UNIPROT:Q38862; GB:AE002093; NID:g4567202; PTDN:AAD23618.1; GSPDB:GN
A:Gene: At2g22240
A:Map position: 2
C:Superfamily: Myo-inositol-1-phosphate synthase

Query Match 90.6%; Score 2384; DB 2; Length 510;
Best Local Similarity 88.2%; Pred. No. 1.7e-152;
Matches 450; Conservative 37; Mismatches 23; Indels 0; Gaps 0;

Qy 1 MFIEFKVESPNVXYTETETESVYVYETTELVHNRNGTYQWIVKPKSVNYQFKTNTHVP 60
Db 1 MFIEFKVESPNVXYTETENISVVDYETTELVHNRNGTYQWIVKPKVKTDPKTDIVP 60
Qy 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANFYGSLTQASAIRVSGFOGEE 120
Db 61 KLGVMLVGMGNGNSTLTAGVIANKEGISWATKDVQOANFYGSLTQASSIRVSGYNEE 120
Qy 121 IYAPFKSLPWNPNDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESVPLPGI 180
Db 121 IYAPFKSLPWNPNDDIVFGGWDISNNMLADAMARAVLIDLOKQLRPYMENMPLPGI 180
Qy 181 YDPDFIAANOEERANNVIGTKQBOVOQIIKDIKAFKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANQGRANSVIGTKKEQVDHIKDMREFKEKNKVDKVVLMTANTERYSNVI 240
Qy 241 VGLNDTMENLLAADRNEAIESPSTLYAIACVMEVNPFGINSPOPTFVPGILDLAIANT 300
Db 241 VGLNDTMENLLASVSKDEISEPSTLYAIACVLEBIPFGINSPOPTFVPGILELAISNC 300
Qy 301 LIIGDDFKSGQTKKKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMLASAPQTFRSKEISKS 360
Db 301 LIIGDDFKSGQTKKKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMLASAPQTFRSKEISKS 360
Qy 361 NVVDDMVNSNAILYEPGHPDHVVVVKVPYVGDSCRAMDEYTSEIFMGKNNTVLHNTC 420
Db 361 NVVDDMVNSNAILYEPGHPDHVVVVKVPYVGDSCRAMDEYTSEIFMGGRNTVLHNTC 420
Qy 421 EDSLAAPIIIDVLLAELSTRIOFKANEGKFSFHPVATILSYLTAKALVPFGTPVYN 480
Db 421 EDSLAAPIIIDVLLAELSTRIOFKABEGKFSFHPVATILSYLTAKALVPFGTPVYN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 3
T50021
inositol-3-phosphate synthase (EC 5.5.1.4) T31P16.160 [similarity] - Arabidopsis thaliana
```

```
N:Alternate names: protein T31P16.160
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T50021
R:Beyan, M.; Zimmermann, W.; Grueneisen, A.; Wandt, R.; Kalicki, J.; Woldmann, P.; Smi
submitted to the Protein Sequence Database, May 2000
A:Reference number: 225027
A:Accession: T50021
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-510 <BEV>
A:Cross-references: UNIPROT:Q9LX12; EMBL:AJ356332; GSPDB:GN00063; ATSP:T31P16.160
A:Experimental source: cultivar Columbia; BAC clone T31P16
C:Genetics:
A:Gene: ATSP:T31P16.160
A:Map position: 5
A:Insertions: 63/2; 86/2; 131/3; 214/2; 290/1; 328/3; 387/3; 450/3; 471/3
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 90.3%; Score 2376; DB 2; Length 510;
Best Local Similarity 87.8%; Pred. No. 5.9e-152;
Matches 448; Conservative 40; Mismatches 22; Indels 0; Gaps 0;

Qy 1 MFIEFKVESPNVXYTETETESVYVYETTELVHNRNGTYQWIVKPKSVNYQFKTNTHVP 60
Db 1 MFIEFKVESPNVXYTETENISVVDYQTELVHNRNGKGAFCQWTKVFKYEFKTDTHVP 60
Qy 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIQOANFYGSLTQASAIRVSGFOGEE 120
Db 61 KLGVMLVGMGNGNSTLTAGVIANREGISWATKDVQOANFYGSLTQASSIRVSGFGE 120
Qy 121 IYAPFKSLPWNPNDDIVFGGWDISNNMLADAMARAKVFDIDLOKQLRPYMESVPLPGI 180
Db 121 IYAPFKSLPWNPNDDIVFGGWDISNNMLADAMARAVLIDLOKQMRPFVEHNVPLPGI 180
Qy 181 YDPDFIAANOEERANNVIGTKQBOVOQIIKDIKAFKATKVDKVVLMTANTERYSNLV 240
Db 181 YDPDFIAANQGRANHVIGTKKQLEQVHIDREFKEKNVDKVVLMTANTERYSNV 240
Qy 241 VGLNDTMENLLAADRNEAIESPSTLYAIACVMEVNPFGINSPOPTFVPGILDLAIANT 300
Db 241 VGLNDTMENLLASSLDKDEAIESPSTLYAIACVLENVFPFGINSPOPTFVPGILELAIRNC 300
Qy 301 LIIGDDFKSGQTKKKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMLASAPQTFRSKEISKS 360
Db 301 LIIGDDFKSGQTKKKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMLASAPQTFRSKEISKS 360
Qy 361 NVVDDMVNSNAILYEPGHPDHVVVVKVPYVGDSCRAMDEYTSEIFMGKNNTVLHNTC 420
Db 361 NVVDDMVNSNAILYEPGHPDHVVVVKVPYVGDSCRAMDEYTSEIFMGKNNTVIMHNTC 420
Qy 421 EDSLAAPIIIDVLLAELSTRIOFKANEGKFSFHPVATILSYLTAKALVPFGTPVYN 480
Db 421 EDSLAAPIIIDVLLAELSTRIOFKANEGKFSFHPVATILSYLTAKALVPFGTPVYN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 4
T08436
inositol-3-phosphate synthase (EC 5.5.1.4) [similarity] - rape
C:Species: Brassica napus (rape)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T08436
R:Hussein, A.; Bourgeois, J.; Polvi, S.; Tsang, E.; Keller, W.A.; Georges, F.
submitted to the EMBL Data Library, August 1996
A:Reference number: 216418
A:Accession: T08436
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
```



```
|||||
Db IYAFKSLIPWNPDDIVFGGWDISNNMLADSMTRAKLIDLIDLOKOLRPYMESWVPLPGI 180
121
181 YDPDFIAANOBERANNVIGTGOBOVOOIIKDIAFKREATVDKVVVMTANTERYSNLV 240
181 YDPDFIAANOGRANNSVIGTKEQVEQIIKDIREFEKKKNVDKIVMTANTERYSNVC 240
241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWMEVPPINGSPONTFVPGIIDLAIARNT 300
241 AGLNDTMENTLASVDKNEAEVSPSTLYAIACWMEVPPINGSPONTFVPGIIDLAIKNC 300
301 LIIGDDFFSGQTKKMSVLDVFLVAGIKPTISIVSYNHLGNNDGNMLSAPQFRSKEISKS 360
301 LIIGDDFFSGQTKKMSVLDVFLVAGIKPTISIVSYNHLGNNDGNMLSAPQFRSKEISKS 360
361 NVVDDMVNSNALIYEGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLHNTC 420
361 NVVDDMVNSNALIYEGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLHNTC 420
421 EDSLAAPIIIDLVLAELSTRIOFKAENEGKFHSPHVAITLSYLTAKAPLVPPGTPVNV 480
421 EDSLAAPIIIDLVLAELSTRIOFKAENEGKFHSPHVAITLSYLTAKAPLVPPGTPVNV 480
481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
481 ALAKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 7

```
T05017
inositol-3-phosphate synthase (EC 5.5.1.4) T19P19.190 [similarity] - Arabidopsis thaliana
N:Alternate names: protein T19P19.190
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05017
R:Bevan, M.; Monfort, A.; Csaacuberta, E.; Puigdemenech, P.; Hohenseil, J.; Mewes, H.W.;
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15394
A:Accession: T05017
A:Molecule type: DNA
A:Residues: 1-511 <BEV>
A:Cross-references: UNIPROT:P42801; EMBL:AL022605
A:Experimental source: cultivar Columbia; BAC clone T19P19
C:Genetics:
A:Map position: 4
A:Introns: 64/2; 132/3; 215/2; 291/1; 329/3; 388/3; 451/3
A>Note: T19P19.190
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD
```

Query Match 89.3%; Score 2349.5; DB 2; Length 511;
Best Local Similarity 87.9%; Pred. No. 3.6e-150;
Matches 449; Conservative 32; Mismatches 29; Indels 1; Gaps 1;

```
1 MFENFKVESPNVXYTETETIQSVNYETTELVHENR-NGTYQWIVPKSVNYQFKTNTHV 59
|||||
1 MFISBFKVESPNVXYTETENIHSHVDYETTELVEHKTNGVQWIVPKSVNYQFKTNTHV 60
60 PRLGMLVGMGNNSSTLTGVIARREDISWATKTKIOQANFSGLTQASAIRVSGFGE 119
61 PRLGMLVGLGNNSSTLTGVIARKEGISWATKTKIOQANFSGLTQASSIRVSGFGE 120
120 EYAPFKSLIPWNPDDIVFGGWDISNNMLADAMARAKVFDIDLQOLRPYMESWVPLPG 179
121 EYAPFKSLIPWNPDDIVFGGWDISNNMLADAMARAKVFDIDLQOLRPYMENIVPLPG 180
180 IYDPDFIAANOBERANNVIGTGOBOVOOIIKDIAFKREATVDKVVVMTANTERYSNLV 239
181 IYDPDFIAANOGRANNSVIGTKEQVEQIIKDIREFEKKKNVDKVVVMTANTERYSNLV 240
240 VGLNDTMENTLAAVDRNEAISPSTLYAIACWMEVPPINGSPONTFVPGIIDLAIARN 299
241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWMEVPPINGSPONTFVPGIIDLAIARN 300
```

```
QY 300 TLIGDDFFSGQTKKMSVLDVFLVAGIKPTISIVSYNHLGNNDGNMLSAPQFRSKEISK 359
Db 301 VLIIGDDFFSGQTKKMSVLDVFLVAGIKPTISIVSYNHLGNNDGNMLSAPQFRSKEISK 360
QY 360 SNVDDMVNSNALIYEGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLHNT 419
Db 361 SNVDDMVNSNALIYEGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLHNT 420
420 CEDSLAAPIIIDLVLAELSTRIOFKAENEGKFHSPHVAITLSYLTAKAPLVPPGTPVNV 479
421 CEDSLAAPIIIDLVLAELSTRIOFKAENEGKFHSPHVAITLSYLTAKAPLVPPGTPVNV 480
480 NALSKORAMLENIMRACVGLAPENNMILEYK 510
481 NALSKORAMLENIMRACVGLAPENNMILEYK 511
```

RESULT 8

```
T04339
inositol-3-phosphate synthase (EC 5.5.1.4) - barley
C:Species: Hordeum vulgare (barley)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T04339
R:Larson, S.R.; Raboy, V.
submitted to the EMBL Data Library, March 1998
A:Description: Linkage mapping maize and barley myo-inositol 1-phosphate synthase genes.
A:Reference number: Z14366
A:Accession: T04339
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-510 <LAR>
A:Cross-references: UNIPROT:O65195; EMBL:AF056325; NID:g3152730; PIDN:AACT1133.1; PID:g31
A:Experimental source: cv. Harrington
C:Genetics:
A:Gene: INO1
A:Map position: 4
A:Function:
A:Description: catalyzes reversible isomerization of D-glucose 6-phosphate to 1L-myo-inoe
A:Pathway: inositol biosynthesis
A>Note: first step
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD
```

Query Match 88.9%; Score 2338; DB 2; Length 510;
Best Local Similarity 88.0%; Pred. No. 2.1e-149;
Matches 449; Conservative 26; Mismatches 35; Indels 0; Gaps 0;

```
1 MFENFKVESPNVXYTETETIQSVNYETTELVHENRNGTYQWIVPKSVNYQFKTNTHV 60
|||||
1 MFISBFKVESPNVXYTETENIHSHVDYETTELVEHSHDGSKVVVRPKSVNYHFKTNTHV 60
61 KLGLMLVGMGNNSSTLTGVIARREDISWATKTKIOQANFSGLTQASAIRVSGFGE 120
61 KLGLMLVGMGNNSSTLTGVIARREDISWATKTKIOQANFSGLTQASAIRVSGFGE 120
121 IYAFKSLIPWNPDDIVFGGWDISNNMLADAMARAKVFDIDLQOLRPYMESWVPLPGI 180
121 IYAFKSLIPWNPDDIVFGGWDISNNMLADAMARAKVFDIDLQOLRPYMESWVPLPGI 180
181 YDPDFIAANOBERANNVIGTGOBOVOOIIKDIAFKREATVDKVVVMTANTERYSNLV 240
181 YDPDFIAANOGRANNSVIGTKEQVEQIIKDIREFEKKKNVDKVVVMTANTERYSNLV 240
241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWMEVPPINGSPONTFVPGIIDLAIARNT 300
241 VGLNDTMENTLAAVDRNEAISPSTLYAIACWMEVPPINGSPONTFVPGIIDLAIKNC 300
301 LIIGDDFFSGQTKKMSVLDVFLVAGIKPTISIVSYNHLGNNDGNMLSAPQFRSKEISKS 360
301 LIIGDDFFSGQTKKMSVLDVFLVAGIKPTISIVSYNHLGNNDGNMLSAPQFRSKEISKS 360
361 NVVDDMVNSNALIYEGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKNTIVLHNTC 420
```


A:Residues: 1-525 <W1>
A:Cross-references: UNIPROT:Q16664; EMBL:AL033553; PDB:CAA22132.1; CESP:VF13D12L.1
A:Experimental source: clone VF13D12L
R:Gajdety, S.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z19209
A:Accession: J20002
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-525 <W12>
A:Cross-references: EMBL:Z69902; PDB:CAA93771.1; GSPDB:GN00020; CESP:VF13D12L.1
A:Experimental source: clone CA7D12
C:Genetics:
A:Gene: CESP:VF13D12L.1
A:Map position: 2
A:introns: 106/2; 287/1; 411/2
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase

Query Match 53.0%; Score 1393.5; DB 2; Length 525;
Best Local Similarity 53.3%, Pred. No. 6.4e-86;
Matches 269; Conservative 93; Mismatches 134; Indels 9; Gaps 5;

Dy VESPNVKTETREISOVVYETTELVEHNRNNTGYQMIVPKSVNQFKTTHVPLGLVMVL 67
||| | :
15 VESPVNKLEDVALSRFFRYKRNHEH-RADGLH--VTPKHDSFTKLKPRKTKTLILV 70
||| | :
68 GWGGNGSTLLGVANREDISWATKDKIQQANYFGSLTQAASIRVG-SFOGEIYAPF 125
||||| :
71 GLGGNGSTANGSIFANGYAMTWRTKEHSQANFYGSVTGTATVAHLGDASTAQNGIVPF 130
||||| :
Dy KSLLPNVNDPDIYFGGWDISMNMTADAMARKVDIDLOKOLRPYMESAVPLPGIYPDF 185
::: :::::
131 KDVPILSPNDLIISGMDISDNLYEAMGRAKVPEPELOEKLRPFMEPIVPLPSIYYPDF 190
||| | :
Dy IAAOGEERANNVIKG-TQEQVOQIKDKIAFKKATKDVKYVMTANTERYSNLVGLN 244
||: ||: ||| | :
186 IASNOGDANNVIPEDNKLLEHHRADIRKFQSHHECVIYMTANTEHYTDVRGGLN 250
||: ||: ||| | :
191 IASNOGDANNVIPEDNKLLEHHRADIRKFQSHHECVIYMTANTEHYTDVRGGLN 250
||: ||: ||| | :
Dy DTMENLAAVDRNEAESPTLYAIACMVENVPPINGSPOTFVPGILDIAIAENTLIG 304
:
251 ATADEIMESIYNVEDEVSPSNI FAVASILEGAHYINSGPQTLLVPGILELRHKVFPVG 310
||| | :
Dy DDFSGSQTKMKSVLYDPLVAGIKPITSIVSYNHGNNDGNMLSAHPQTFRSKEISKSVND 364
||| | :
311 DDFFSGSQTKMKSAFDFLVSSGMRKPESIVSYNHGNNDGNMLSEAROFRSKEISKSVND 370
||| | :
Dy DMVNSMNLIEPGEHPDHVVVKYVPVYGSKRAMDEXTSII FMGKNRTVLTHTCCDSL 424
||| | :
371 DMVNSMNLIEPDANKPDICVVIKVPYVADSKRAMDETCSI FMGKQTFVFNHTCCDSL 430
||| | :
Dy LAAPILIIDLVLAEISTRIOPKAENEGKFHFPHPATILSYLTAPLVPCTGPVNALSK 484
||| | :
431 LASPLIYLALITELASVSXKVDDE--YRFHFVSLTSLSLRLARPVPBPOTPIISNFMR 488
||| | :
Dy QPAMLENIMRAQCVGLAPENNMLEY 509
||| | :
489 QESTLKLVLTALAGPPSDTDWQIER 513
||| | :

RESULT 12
S45452 inositol-3-phosphate synthase (EC 5.5.1.4) - yeast (Candida albicans)
C:Species: Candida albicans
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S45452
R:Kilig, L.S.; Zobel, P.A.; Devry, C.G.; Losberger, C.
A>Title: Comparison of INOI gene sequences and products in Candida albicans and Saccharomyces cerevisiae
A:Reference number: S45452; MUID:95066381; PMID:7975896
A:Accession: S45452
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA

A:Residues: 1-520 <K1>
A:Cross-references: UNIPROT:P42800; EMBL:L22737; NID:g413750; PDB:AAA62849.1; PID:g69575
C:Genetics:
A:Gene: INO1
C:Function:
A:Description: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo-
A>Note: requires NAD
C:Superfamily: myo-inositol-1-phosphate synthase
C:Keywords: intramolecular lyase; isomerase; NAD

Query Match 52.8%; Score 1389.5; DB 2; Length 520;
Best Local Similarity 52.1%; Pred. No. 1.2e-85;
Matches 267; Conservative 98; Mismatches 132; Indels 15; Gaps 5;

QY 9 ESPNVKKTETELQSYVYNTETTELVAHNRNGTQWIKPKSVNQFKNTHVPLGLVWLVG 68
Db 8 KSSNVSTVDHLYKTFEYENS-VVERKDANG--KFIVTPASDIEFFKVDLPPVPGVLLLVG 64
QY 69 WGNNGNSTLTGTVIANREDISWATKDKIOQANVPFGLTQASAIRVG--SPQGEIYAPFK 126
Db 65 IGGNNGITLLGATLTKANKHNSIFENKEGVKPKYGSVTOASTYKIGVDKETGSDVYVFN 124
QY 127 SLLEPVNPDIIYFGGMDISNNMLADAMARAKYFDIDLQKQRLPYMESMVPPLPGIYDPDFI 186
Db 125 SIPEPVNPNNDLVVQDMDISGLPLDQAMKRAKVLDTLQKQVLYLENKKPLRESIYVDFI 184
QY 187 AANQEEANNAVYK-----GTKQEQVQIITKDIKAFKFAKTVKDVVVLMTANTERYSN 238
Db 185 ALNQSERRANNVFNQVNGEVKTDNKNADYEKIKRDIRDFAKKELDKVITILMTANTERYAD 244
QY 239 LVVGLNDPMENLLAVDNEAEISPSLLYAIQWENVPFINGSPONTFVPGILDLAIAR 298
Db 245 VLPVNDTPADNLIKSIKSHSHEIAPSTVPAVASILEKVPYINGSPONTFVPGVIELAEKY 304
QY 299 NTLIGDDPFKSGQTKAKSVLVDPLVAGAGIKPISIVSYNHLGNNDGNMLSPQTFRSKEIS 358
Db 305 DSFIGDDPFKSGQTKIKSVLAQFLVDAGIKPLISINSYNHLGNNDGNVLSPPQFRSKEIS 364
QY 359 KSNVVDVMSNALLY--EPGEHPHVVIKTVPYVGSKRPMDEYTSSEIFMGKNTIYL 416
Db 365 KQSVVDIDIESENELLYKNKSGDKVDHCIVIKYLPVAGDSKVMDEYVSELMGLGNKIS 424
QY 417 HNTGDSLLAPPIILDVLLAELSTRIOFKANEKGHFSFHPATLSYLTKAPLVPGT 476
Db 425 HNVGDSLLAPPLIIDLVAVEFATRVQVKGKSDYDELVPVASLSTLYLKAPDLARPGF 484
QY 477 PVVNAISKQAMLENIMRACVGLAPENNMLE 508
Db 485 KPINGANKQKQQLVNLISLVGLPIDNELRFE 516

RESULT 13
A30902
Inositol-3-phosphate synthase (EC 5.5.1.4) [validated] - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein U0610; protein YOL153C
C:Species: Saccharomyces cerevisiae
C:Date: 18-Apr-1988 #sequence revision 08-Sep-1995 #ext change 16-Aug-2004
C:Accession: S55160; B32209; S56935; S71644; A30877; A30902
R:Karsoulou, C.; Tzermila, M.; Alexandraki, D.
A:Description: The EMBL Data Library, May 1995
Submitted to the complete sequence of a 40.7 kb segment located on the left arm of yeast
yeast hypothetical proteins.
A:Reference number: S55159
A:Accession: S55160
A:Molecule type: DNA
A:Residues: 1-555 <KAT>
A:Cross-references: EMBL:X87371; NID:g854542; PID:g854544
R:Dean-Johnson, M.; Henry, S.A.
J. Biol. Chem. 264, 1274-1283, 1989
A:Title: Biosynthesis of inositol in yeast. Primary structure of myo-inositol-1-phosphate
A:Reference number: A32209; MUID:89093118; PMID:2642502
A:Accession: B32209
A:Molecule type: DNA

A:Residues: 23-35, 'RL', 37-81, 'FE', 83-87, 'TRNVAHWVW', 88, 'QQW', 92-103, 'WPRYRISTMTWS', 116-
VISFORLISFSAAYL, <DEA>
A:Cross-references: EMBL:J04453
R:Katsoulou, C.; Tsermli, M.; Alexandraki, D.
Submitted to the Protein Sequence Database, September 1995
A:Reference number: S56912
A:Accession: S56935
A:Molecule type: DNA
A:Residues: 1-555 <KAW>
A:Cross-references: EMBL:Z49428; NID:g1015570; PID:g1015571; MIPS:YJL153c
R:Katsoulou, C.; Tsermli, M.; Tavernarakis, N.; Alexandraki, D.
Yeast 12, 787-797, 1996
A:Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome X
of chromosome XI.
A:Reference number: S71643; MUID:96408771; PMID:8813765
A:Accession: S71644
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-555 <KAP>
A:Cross-references: EMBL:X87371; NID:g854542; PID:CAA6802.1; PID:g854544
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995
A:Gene: SGD:INO1
A:Cross-references: SGD:S0003689; MIPS:YJL153c
A:Map position: 10L
C:Complex: homotetramer
C:Function: catalyzes the reversible isomerization of D-glucose 6-phosphate to 1L-myo
A:Pathway: inositol biosynthesis
A:Note: requires NAD
C:Superfamily: Myo-inositol-1-phosphate synthase
C:Keywords: cytosol; homotetramer; intramolecular lyase; isomerase; NAD

Query Match 52.8%; Score 1389.5; DB 2; Length 555;
Best Local Similarity 53.1%; Pred. No. 1.3e-85;
Matches 278; Conservative 87; Mismatches 138; Indels 21; Gaps 8;

QY 3 IENFVSEPNKYTETETIQSVYNYETTELHNRNGTQWIKPKSVNYQFNTNHP-K 61
DB 31 ITSVAVVDKCTYKNELETLTKYSYENA-VYTKTASGRD-VTPVQDVDFLDDKKBEK 87
QY 62 LGVNLVWGCGNNGSTLTGCVIANREDISWATKDKIQANFYGSLTQASAIRVG-SFOGEE 120
DB 88 LGIMLIGCGNNGSTLTVASVLANKHNEFQTEGKQRPYFBSMTQCSTLKIGIABGND 147
QY 121 IYAPFKSLPMVNPDDIVFGGNDISNMNLADMAKAFYDIDLOKQLEPYMESWPLRGI 180
DB 148 IYAPFNLSLPMVSPMDPVVSGMDINNADLYEAMQSRQVLEVDLOQLKAKMSLVKPLPSI 207
QY 181 YDPDPIAANOEBRANV-----KGT-----KQOVQOIIKDIKAKREATKDKVYVMTA 231
DB 208 YYPDPIAANOEBRANNCINLDEKGNVYTRGKWTNLIQIRIRDIQNFKEBNALDKVILVMTA 267
QY 232 NTERSNLWGLNDMENLILAAVDNBEAISPSTLYAIAQWENYPTNGSPONTFVPGI 291
DB 268 NTERKVERSPGVNDMENLQSIKIDHEIASSTIFAAASILSGVPTNGSPONTFVPGI 327
QY 292 IDLAIAARNTLIGDDPKSGQTKMKSVLVDFVAGIKPTSTIYSYNHLGNDGMSLAPOT 351
DB 328 VOLAHEGTFIAGDDLKSGQTKLSVLAQFLVDAGIKPTSTIASYNHLGNDGMSYNLSAKQ 387
QY 352 FRSKKISKSNVVDVDMVNSNAILYER-GEHPHVIVYIKVYVGDGSKRAMDEYTSIEFMG 409
DB 388 FRSKKISKSNVYIDLIASNDILYNDKLGKVDHCIVIKMKVGVGSKRAMDEYSELIMG 447
QY 410 GKNITVLNTECEDSLAARPIILVLAELSTRIOFK-----AENEGKPHSPHAPATILS 464
DB 448 GHNRSIHNVCEDSLALATPLIILDLVMTBFCRVSYYKAVDPKEDAGKFEENFPVLTLS 507
QY 465 YLTKAIVLPFGTPVNVNALSQKAPAMLEIMRACVGLAPENNMLE 508
DB 508 YMLKAPLTPRPGFHPVNGINKQRTALENFLRLILGIPQNEILAFE 551

RESULT 14
146317
hypothetical protein DKFZp434A0612.1 - human
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46317
R:Ducresthoef, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
Submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23035
A:Accession: T46317
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-430 <AAA>
A:Cross-references: UNIPROT:Q9NSU0; EMBL:AL137749
A:Experimental source: adult testis; clone DKFZp434A0612
C:Genetics:
A:Note: DKFZp434A0612.1
C:Superfamily: myo-inositol-1-phosphate synthase

Query Match 49.9%; Score 1312.5; DB 2; Length 430;
Best Local Similarity 64.2%; Pred. No. 1.3e-80;
Matches 244; Conservative 66; Mismatches 69; Indels 1; Gaps 1;

QY 131 MWNPDIVFGGMDISNMNLADAMAKAFYDIDLOKQLEPYMESWPLRGIYDPDPIANO 190
DB 1 MVAEPDVLFDGWDISSILNLAEMRRKAVLDWGLQPLPHEALRPRSVYIPEFIANQ 60
QY 191 EERANNVKTGKQEQVOQIINDIKAFKATKDKVYVMTANTERYSNLVVGLNTMENT 250
DB 61 SARBDNLPGRSAQGLEIRDIRDFRSSAGLDKIVYVMTANTERFCEVIRPDLNANTAL 120
QY 251 LAVDNRNBEAISPSTLYAIAQWENYPTNGSPONTFVPGIIDLAIARNTLIGDDPKSG 310
DB 121 LRTTELGL-LEVPSTLFAVASILBEGCAFINGSPQTLVPGALIELAMQHRVFGGDDPKSG 179
QY 311 QTKMSVYVDFLVAGIRPTSTIVSYNHLGNDGMSAPQFRSKISKSNVVDVDMVSN 370
DB 180 QTKMSVYVDFLVAGIRPTSTIVSYNHLGNDGMSAPQFRSKISKSNVVDVDMVSN 239
QY 371 AILVEGEHPHVIVYIKVYVGDGSKRAMDEYTSIEFMGKNTIYVNTCEPDSLAAPI 430
DB 240 PVLTYRGESEPHCVIKVYVGDGSKRAMDEYTSIEFMGKNTIYVNTCEPDSLAAPI 299
QY 431 IDLVLAELSTRIOFKAENEGKPHSPHAPATILSYTAPLVPFGTPVNVNALSQKAPAMLE 490
DB 300 IDLALITELCQRFVFCFTMDPEPQTFHPVLSLSFLFAPLVPFGTPVNVNALSQKAPAMLE 359
QY 491 NIMRACVGLAPENNMLEYK 510
DB 360 NILRACVGLAPENNMLEHK 379

RESULT 15
T36191
Probable myo-inositol phosphate synthase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36191
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.J.
Submitted to the EMBL Data Library, March 1999
A:Reference number: Z21600
A:Accession: T36191
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-388 <SAU>
A:Cross-references: UNIPROT:Q9Z4Y6; EMBL:AL035707; PID:CA838867.1; GSPDB:GN00070; SCOECD
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOECD:SC829.12c
C:Superfamily: myo-inositol-1-phosphate synthase

Query Match 11.1%; Score 291; DB 2; Length 388;

Best Local Similarity 27.7%; Pred. No. 4,8e-12;
Matches 110; Conservative 61; Mismatches 158; Indels 68; Gaps 11;

```

QY      61 KLGWLVWGMGNNNGSTLUGVIANBEDI:SWAFKDKI:QANVYSGSLTQASAI:VNGSPQGE 120
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2 RTGIMLVWARGSVATTIVVGAALBAGLVPAT-----GCVSALEA-----PDG-- 44

QY      121 IYAPFKSLIPMNPPDIVFGGWDISNNUL--ADAMARAKVPEIDLQOKLRPYMESMBV- 176
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      45 -----VPLPGPELWFGGHDVVGTLVKRAEQLAEGV-----VPR 80

QY      177 -LPGIYDPDFITAAOEEPRANNVIKGTQO-----EVOQOI:IKDKAFKATKYDKV 226
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      81 GLPGVLTZAEI:DAADA:EI:RP--APPGTEEGGDEGTPQASAGAIVADLTGFRRELTGIDRV 138

QY      227 VLMATNTERYSULVVGGLNDYMENTL:LAADNRNEA:ISPTLYAIACMEVNPETNSPQ-N 285
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      139 VVNVSS:TOPPA-VPHPAH:ASAA:LEAL:ERGR:PLPVSS:LYVAAL:RACGA:VVD:FTPS:G 197

QY      286 TFEVGLID:IAI:ARN:LI:IGDDPFKSGQTKWKS:LYVDFLWAG:IKPT:SI:SYNH:GNND:GN 345
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      198 AR:LPAL:DEL:AR:EG:GLP:PYAGSDGK:TBET:LVK:SV:LA:MPFAR:AR:RV:SW:SGT:NL:GGD:GAT 257

QY      346 L:SA:PO:FSK:KEI:SKSNVVD:VWNSA:NL:YE:PE:GH:PDH:VV:IKV:PYV:GDSK:AMDE:Y:SE 405
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      258 LADPERVVS:KNAS:KGL:VLE-----AELGHA:VEG:GVI:HHV:DLGEM:KTA:MDH:V:PE 308

QY      406 IFMGKNTI:VLAHNTCED:SLA:PI:ILD:VILAE:LS:R 442
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      309 GFLGAR:MTLQ:FTWGC:DSSLA:PLV:LDLAR:PMAL:LR 345

```

Search completed: June 7, 2005, 16:47:42
Job time : 26.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 15:09:19 ; Search time 92.5 Seconds
(without alignments)
2132.409 Million cell updates/sec

Title: US-10-718-952-12

Perfect score: 2631
Sequence: 1 MFENFKVESPNVKTETETI.....NIMRACVGLAPENNILSYK 510

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 366760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneeqp19808:*\n2: geneeqp19908:*\n3: geneeqp20008:*\n4: geneeqp20018:*\n5: geneeqp20028:*\n6: geneeqp20038:*\n7: geneeqp20038:*\n8: geneeqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2631	100.0	510	ADQ14501	Adq14501 Mutant bo
2	2631	100.0	510	AD882002	Ad882002 Soybean m
3	2624	99.7	510	ADQ14503	Adq14503 Mutant bo
4	2624	99.7	510	ADQ14505	Adq14505 Wild type
5	2624	99.7	510	AD882004	Ad882004 Soybean m
6	2624	99.7	510	AD882006	Ad882006 Soybean m
7	2593	98.6	510	AAW79740	AAw79740 Soybean w
8	2593	98.6	510	ADQ14499	Adq14499 Mutant bo
9	2593	98.6	510	ADQ14491	Adq14491 Wild type
10	2593	98.6	510	AD882000	Ad882000 Soybean m
11	2593	98.6	510	AD881994	Ad881994 Soybean m
12	2588	98.4	510	AAW79741	AAw79741 Soybean m
13	2588	98.4	510	ADQ14495	Adq14495 Mutant bo
14	2588	98.4	510	AD881998	Ad881998 Soybean m
15	2470	93.9	536	AAW24477	AAw24477 Nicotiana
16	2431	92.4	505	ADG73738	Adg73738 Myo-inosi
17	2384	90.6	510	AAAG50575	AAg50575 Arabidops
18	2384	90.6	510	ADW73525	ADw73525 Thale cre
19	2384	90.6	581	AAAG50574	AAg50574 Arabidops
20	2384	90.6	645	AAAG50573	AAg50573 Arabidops
21	2376	90.3	510	AAAB48935	AAb48935 Brassica
22	2354	89.3	510	AAW96259	AAw96259 Phytate p
23	2349.5	89.3	511	AAAG09861	AAg09861 Arabidops
24	2349.5	89.3	511	AAAG32501	AAg32501 Arabidops
25	2349.5	89.3	534	AAAG09860	AAg09860 Arabidops

26	2346	89.2	510	2	AAW97882	AAw97882 Maize myo
27	2345	89.1	510	8	ADP43920	Adp43920 Rice L-my
28	2337	88.8	510	4	AAAB47286	AAb47286 MIP synth
29	2084	79.2	446	3	AAAG32502	AAg32502 Arabidops
30	2084	79.2	446	3	AAAG09862	AAg09862 Arabidops
31	1883	71.6	512	8	ADP43919	ADp43919 Porterei
32	1779	67.6	380	3	AAAG32503	AAg32503 Arabidops
33	1616	61.4	558	4	AAAG65577	AAg65577 Human hml
34	1616	61.4	558	4	AAAM40076	AAm40076 Human pol
35	1616	61.4	558	4	AAAB92849	AAb92849 Human pro
36	1616	61.4	558	4	AAAB93732	AAb93732 Human pro
37	1616	61.4	558	5	AAO18891	AAo18891 Human ova
38	1612.5	61.3	534	6	ABJ25857	ABj25857 Aspergill
39	1608.5	61.1	534	6	ABJ26457	ABj26457 Aspergill
40	1602	60.9	557	4	AAAG65576	AAg65576 Murine m
41	1602	60.9	565	4	ABAB65497	ABb65497 Drosophi
42	1565.5	59.5	659	8	ADH99714	ADh99714 Novel hum
43	1538	58.5	505	8	ADG73737	ADg73737 Aspergill
44	1538	58.5	532	8	ADG73736	ADg73736 Aspergill
45	1429	54.3	504	7	ADB64735	ADb64735 Human pro

ALIGNMENTS

RESULT 1		
ID	ADQ14501	ADQ14501 standard; protein; 510 AA.
XX		
AC	ADQ14501;	
XX		
DT	23-SEP-2004	(first entry)
XX		
DE	Mutant soybean myo-inositol 1-phosphate synthase polypeptide #3.	
XX		
KW	Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;	
KM	raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;	
KW	mutant; mulein.	
XX		
OS	Glycine max.	
OS	Synthetic.	
XX		
PN	US2004128713-A1.	
XX		
PD	01-JUL-2004.	
XX		
PF	21-NOV-2003; 2003US-00718952.	
XX		
PR	08-APR-1997; 97US-00835751.	
PR	07-APR-1998; 98WO-US006822.	
PR	26-APR-1999; 99US-0029315.	
PR	11-MAR-2002; 2002US-00025003.	
XX		
PA	(HITZ/) HITZ W D.	
PA	(SEBA/) SEBASTIAN S A.	
PA	(GRAC/) GRACE D J.	
PA	(STRE/) STREIT L G.	
XX		
PI	Hitz WD, Sebastian SA, Grace DJ, Streit LG;	
XX		
DR	WPI; 2004-533135/51.	
DR	N-PSDB; ADQ14500.	
XX		
PT	New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,	
PT	useful for producing plants with decreased raffinose, stachyose, and	
PT	phytic acid and increased sucrose, leading to valuable and useful soybean	
PT	products.	
XX		
PS	Claim 9; SEQ ID NO 12; 48pp; English.	
XX		
CC	The invention relates to a nucleic acid fragment encoding a soybean myo-	
CC	inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate	
CC	synthase having a decreased capacity for the synthesis of myo-inositol 1-	

Db 61 KLGWLVGWMGNGNSTLTGVIANREDISWATKXIQOANYFGSLTQASAIRVSGFQEE 120
 QY 121 IYAPFKSLIPWNPDDIVFGGWDISNNMLADAMAAKVFDDLOKQARPYMESWVPLGI 180
 Db 121 IYAPFKSLIPWNPDDIVFGGWDISNNMLADAMAAKVFDDLOKQARPYMESWVPLGI 180
 QY 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVVMTANTERYSNLV 240
 Db 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVVMTANTERYSNLV 240
 QY 241 VGLNDTMENLLAAVDNRNABISPSSTLYAIACWENVPPINGSPOPTFVPGGLIDLAIAANT 300
 Db 241 VGLNDTMENLLAAVDNRNABISPSSTLYAIACWENVPPINGSPOPTFVPGGLIDLAIAANT 300
 QY 301 LIGGDDFKSGQTKKMSVVDVFLVNGAGIKPTSVSYNHGNDGNMLSAPOTFRSKEISKS 360
 Db 301 LIGGDDFKSGQTKKMSVVDVFLVNGAGIKPTSVSYNHGNDGNMLSAPOTFRSKEISKS 360
 QY 361 NVVDDMVNSNAILYEPGEHPDHVVVVKVPPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
 Db 361 NVVDDMVNSNAILYEPGEHPDHVVVVKVPPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420
 QY 421 EDSLILAAPIIIDLVLLAELSTRIOFKAENEGKFSFHPVATILSYLTRAPLVPPTPVVN 480
 Db 421 EDSLILAAPIIIDLVLLAELSTRIOFKAENEGKFSFHPVATILSYLTRAPLVPPTPVVN 480
 QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
 Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 3

ID ADO14503 standard; protein; 510 AA.

AC ADO14503;

DT 23-SEP-2004 (first entry)

Db Mutant soybean myo-inositol 1-phosphate synthase polypeptide #4.

XX Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
 KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
 KM mutant; mletin.

OS Glycine max.
 OS Synthetic.

XX US2004128713-A1.

PN 01-JUL-2004.

XX 21-NOV-2003; 2003US-00718952.

PF 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.

PR 26-APR-1999; 99US-00293315.

PR 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ M D.

PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.

XX (STRE/) STREIT L G.

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

XX MPI; 2004-533135/51.
 DR N-PSDB; ADO14502.
 XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytic acid and increased sucrose, leading to valuable and useful soybean
 PT products.

XX Example 8; SEQ ID NO 14; 48bp; English.

PS The invention relates to a nucleic acid fragment encoding a soybean myo-
 XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
 CC phosphate. The invention also relates to a chimeric gene operably linked
 CC to suitable regulatory sequences, where expression of the chimeric gene
 CC results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
 CC comprising the chimeric gene, a method of making the soybean plant, a
 CC seed of the soybean plant, a soy protein product derived from the
 CC processing of soybean seeds, a method of making or producing a soy
 CC protein product and a method of using a soybean plant homozygous for at
 CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 CC nucleic acid fragment and methods are useful for producing plants with
 CC decreased raffinose, stachyose and phytic acid content and increased
 CC sucrose and inorganic phosphate content, leading to valuable and useful
 CC soybean products. This sequence represents a mutant soybean myo-inositol
 CC 1-phosphate synthase polypeptide of the invention.

CC Sequence 510 AA;

Query Match 99.7%; Score 2624; DB 8; Length 510;
 Beet Local Similarity 99.8%; Pred. No. 3.4e-215;
 Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFIEFKVESPNVXYTETETIQSVNYVETTELVHNRNCTYQMVKPSVNYQKTNTHVP 60

Db 1 MFIEFKVESPNVXYTETETIQSVNYVETTELVHNRNCTYQMVKPSVNYQKTNTHVP 60

QY 61 KLGWLVGWMGNGNSTLTGVIANREDISWATKXIQOANYFGSLTQASAIRVSGFQEE 120

Db 61 KLGWLVGWMGNGNSTLTGVIANREDISWATKXIQOANYFGSLTQASAIRVSGFQEE 120

QY 121 IYAPFKSLIPWNPDDIVFGGWDISNNMLADAMAAKVFDDLOKQARPYMESWVPLGI 180

Db 121 IYAPFKSLIPWNPDDIVFGGWDISNNMLADAMAAKVFDDLOKQARPYMESWVPLGI 180

QY 241 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVVMTANTERYSNLV 240

Db 241 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVVMTANTERYSNLV 240

QY 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVVMTANTERYSNLV 240

Db 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAFKKATKVDKVVVMTANTERYSNLV 240

QY 241 VGLNDTMENLLAAVDNRNABISPSSTLYAIACWENVPPINGSPOPTFVPGGLIDLAIAANT 300

Db 241 VGLNDTMENLLAAVDNRNABISPSSTLYAIACWENVPPINGSPOPTFVPGGLIDLAIAANT 300

QY 301 LIGGDDFKSGQTKKMSVVDVFLVNGAGIKPTSVSYNHGNDGNMLSAPOTFRSKEISKS 360

Db 301 LIGGDDFKSGQTKKMSVVDVFLVNGAGIKPTSVSYNHGNDGNMLSAPOTFRSKEISKS 360

QY 361 NVVDDMVNSNAILYEPGEHPDHVVVVKVPPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420

Db 361 NVVDDMVNSNAILYEPGEHPDHVVVVKVPPYVGSKRAMDEYTSSEIFMGKNTIVLHNTC 420

QY 421 EDSLILAAPIIIDLVLLAELSTRIOFKAENEGKFSFHPVATILSYLTRAPLVPPTPVVN 480

Db 421 EDSLILAAPIIIDLVLLAELSTRIOFKAENEGKFSFHPVATILSYLTRAPLVPPTPVVN 480

QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 4

ID ADO14505 standard; protein; 510 AA.

AC ADO14505;

DT 23-SEP-2004 (first entry)

DE Wild type soybean myo-inositol 1-phosphate synthase polypeptide #2.
 XX
 KM Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
 KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.
 OS Glycine max.
 XX US2004128713-A1.
 XX
 XX 01-JUL-2004.
 PD
 XX 21-NOV-2003; 2003US-00718952.
 PF
 XX 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 PR 26-APR-1999; 99US-00299315.
 PR 11-MAR-2002; 2002US-00025003.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PA (GRAC/) GRACE D J.
 PA (STRE/) STREIT L G.
 PI Hiltz WD, Sebastian SA, Grace DJ, Streit LG;
 PI
 XX WPI; 2004-53135/51.
 DR N-PSDB; ADQ14504.
 DR
 PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytic acid and increased sucrose, leading to valuable and useful soybean
 PT products.
 XX
 PS Claim 3; SEQ ID NO 16; 48pp; English.
 XX
 CC The invention relates to a nucleic acid fragment encoding a soybean myo-
 CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
 CC phosphate. The invention also relates to a chimeric gene operably linked
 CC to suitable regulatory sequences, where expression of the chimeric gene
 CC results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
 CC comprising the chimeric gene, a method of making the soybean plant, a
 CC seed of the soybean plant, a soy protein product derived from the
 CC processing of soybean seeds, a method of making or producing a soy
 CC protein product and a method of using a soybean plant homozygous for at
 CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 CC nucleic acid fragment and methods are useful for producing plants with
 CC decreased raffinose, stachyose and phytic acid content and increased
 CC sucrose and inorganic phosphate content, leading to valuable and useful
 CC soybean products. This sequence represents a wild type soybean myo-
 CC inositol 1-phosphate synthase polypeptide of the invention.
 CC
 XX Sequence 510 AA;
 SQ
 Query Match 99.7%; Score 2624; DB 8; Length 510;
 Best Local Similarity 99.8%; Pred. No. 3.4e-215;
 Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 181 YDPPFIANOEBRANNVIKGTQEOVOQI IKD IKAFKEATKVKVVLMTANTERYSNLV 240
 QY 241 VGLNDTMENTLLAADRNEAISPSTLYAIACWENVPFINGSPONTFVPGIDIAIARNT 300
 DB 241 VGLNDTMENTLLAADRNEAISPSTLYAIACWENVPFINGSPONTFVPGIDIAIARNT 300
 QY 301 LIGGDDPKSGQTKKSKVLVDFLVGAGIKPTSIYSYNHLGNNDGNLSAPQTFRSKEISKS 360
 DB 301 LIGGDDPKSGQTKKSKVLVDFLVGAGIKPTSIYSYNHLGNNDGNLSAPQTFRSKEISKS 360
 QY 361 NVVDDMNSNAIILYEPBEHPDHVVITKVPYVGSKRAMEYTSIEFMGKNTVLHNTC 420
 DB 361 NVVDDMNSNAIILYEPBEHPDHVVITKVPYVGSKRAMEYTSIEFMGKNTVLHNTC 420
 QY 421 EDSLAAPIIIDVLALAESTRIOFKAENEGKPHSPHVAITILSYLTKAPLVPGTPVNV 480
 DB 421 EDSLAAPIIIDVLALAESTRIOFKAENEGKPHSPHVAITILSYLTKAPLVPGTPVNV 480
 QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
 DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
 RESULT 5
 ADS82004
 ID ADS82004 standard; protein, 510 AA.
 XX
 AC ADS82004;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Soybean myo-inositol 1-phosphate synthase wild-type 4.
 XX
 KW Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
 KW raffinose; stachyose; sucrose; inorganic phosphate; flaculence.
 XX
 OS Glycine max; line 29018JP03.
 XX
 PN US2003074685-A1.
 PD 17-APR-2003.
 PD
 PF 11-MAR-2002; 2002US-00025003.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PI Hiltz WD, Sebastian SA;
 XX
 XX WPI; 2004-639957/62.
 DR N-PSDB; ADS82003.
 DR
 PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
 PT phytic acid and inorganic phosphate content of soybean seeds.
 PT
 XX
 XX Example 8; SEQ ID NO 14; 34pp; English.
 PS
 XX The invention relates to an isolated nucleic acid fragment encoding a
 XX soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 XX phosphate synthase having decreasing capacity for the synthesis for myo-
 XX inositol 1-phosphate. Also included are a chimeric gene (comprising the
 XX nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 XX complement, subfragment or the complement of the subfragment, operably
 XX linked to suitable regulatory sequences, where expression of the chimeric
 XX gene results in a decrease in expression of an endogenous or native gene
 XX encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 XX comprising the chimeric gene (with a heritable phenotype of a seed phytic
 XX acid content of less than 17 micromol/g, a seed content of raffinose plus

CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
 CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype
 CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homologous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with
 CC heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phylic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence represents a wild-type myo-
 CC inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 99.7%; Score 2624; DB 8; Length 510;

Best Local Similarity 99.8%; Pred. No. 3.4e-215;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MFIENFKYSPVVKKTETETIQTQSVNYETTELVEHNRNGTYOMIYKPKSVNYQFKNTHTVP 60
 DB 1 MFIENFKYSPVVKKTETETIQTQSVNYETTELVEHNRNGTYOMIYKPKSVNYQFKNTHTVP 60
 OY 61 KLGVVLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
 DB 61 KLGVVLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
 OY 121 IYAPFKSLLPVNPDPDIVFGGWDISNMNLADAMARAKYFDIDLQKQLRPYMESWYPLDGI 180
 DB 121 IYAPFKSLLPVNPDPDIVFGGWDISNMNLADAMARAKYFDIDLQKQLRPYMESWYPLDGI 180
 OY 181 YDPPDIANOEERANNVIKGTQEOVOQI IKDKIAFKKATKVDKVVVMTANTERYSNLY 240
 DB 181 YDPPDIANOEERANNVIKGTQEOVOQI IKDKIAFKKATKVDKVVVMTANTERYSNLY 240
 OY 241 VGLNDTMENLAAVVRNEAEISPTLYAIACVMEVPIFNGSPONTFVPGILDLAIARNT 300
 DB 241 VGLNDTMENLAAVVRNEAEISPTLYAIACVMEVPIFNGSPONTFVPGILDLAIARNT 300
 OY 301 LIGGDDFKSGQTKMSVLDVFLVGAIGIKPTSIIVSYNHLGNDGMNLASAPQTFRSKEISKS 360
 DB 301 LIGGDDFKSGQTKMSVLDVFLVGAIGIKPTSIIVSYNHLGNDGMNLASAPQTFRSKEISKS 360
 OY 361 NVVDPMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKNTIVLAHTC 420
 DB 361 NVVDPMVNSNALIYEPGEHPDHVVVIKYVPYVGSKRAMEYTSSEIFMGKNTIVLAHTC 420
 OY 421 EDSLILAAPIILDLVLAELSTRIOFKANEGKEHSFHPATILSLVLTAPLVPPTGPVNV 480
 DB 421 EDSLILAAPIILDLVLAELSTRIOFKANEGKEHSFHPATILSLVLTAPLVPPTGPVNV 480
 OY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
 DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 6

ADSS82006 ADSS82006 standard; protein; 510 AA.

XX AC ADSS82006;

XX DT 18-NOV-2004 (first entry)

XX DE Soybean myo-inositol 1-phosphate synthase wild-type 2.

KW Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phylic acid;
 KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.

XX Glycine max; cultivar Wye.

XX US2003074685-A1.

XX 17-APR-2003.

XX 11-MAR-2002; 2002US-00025003.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US0006822.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX Hitz WD, Sebastian SA;

XX WPI: 2004-639957/62.

XX N-PSDB; ADSS82005.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 XX phosphate synthase, useful for altering raffinose saccharide, sucrose,
 XX phylic acid and inorganic phosphate content of soybean seeds.

XX Claim 3; SEQ ID NO 16; 34bp; English.

CC The invention relates to an isolated nucleic acid fragment encoding a
 CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 CC phosphate synthase having decreasing capacity for the synthesis for myo-
 CC inositol 1-phosphate. Also included are a chimeric gene (comprising the
 CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 CC complement, subfragment or the complement of the subfragment, operably
 CC linked to suitable regulatory sequences, where expression of the chimeric
 CC gene results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 CC comprising the chimeric gene (with a heritable phenotype of a seed phylic
 CC acid content of less than 17 micromol/g, a seed content of raffinose plus
 CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
 CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype
 CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homologous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with
 CC heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phylic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence represents a wild-type myo-
 CC inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 99.7%; Score 2624; DB 8; Length 510;

Best Local Similarity 99.8%; Pred. No. 3.4e-215;

Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MFIENFKYSPVVKKTETETIQTQSVNYETTELVEHNRNGTYOMIYKPKSVNYQFKNTHTVP 60
 DB 1 MFIENFKYSPVVKKTETETIQTQSVNYETTELVEHNRNGTYOMIYKPKSVNYQFKNTHTVP 60
 OY 61 KLGVVLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
 DB 61 KLGVVLVGMGNGNSTLTGVIANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120

```
QY 121 IYAPFKSLIPVNPDDIVFGGMDISNMNLADAMARAVFDIDLQKOLRPYMESWVPLPGI 180
DB 121 IYAPFKSLIPVNPDDIVFGGMDISNMNLADAMARAVFDIDLQKOLRPYMESWVPLPGI 180
QY 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAKFKEATKVDKVVVMTANTERSYSLV 240
DB 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAKFKEATKVDKVVVMTANTERSYSLV 240
QY 241 VGLNDTMENLLAAVDNRNEAISPSTLYAIACMVENVPPINGSPONTFVPGILDLAIANT 300
DB 241 VGLNDTMENLLAAVDNRNEAISPSTLYAIACMVENVPPINGSPONTFVPGILDLAIANT 300
QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGNMLAPQFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGNMLAPQFRSKEISKS 360
QY 361 NVDDDMNNSNAILKEPGEHPHVIVIKVPPYVGSKRAMDEYTSSEIFMGKNITVLHNTC 420
DB 361 NVDDDMNNSNAILKEPGEHPHVIVIKVPPYVGSKRAMDEYTSSEIFMGKNITVLHNTC 420
QY 421 EDSLILAAPIIIDLVLALSTRIOFKANEKGKFSFHPVATILSYLTKAPLVPPGTPVNV 480
DB 421 EDSLILAAPIIIDLVLALSTRIOFKANEKGKFSFHPVATILSYLTKAPLVPPGTPVNV 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 7

```
AAW79740
ID AAW79740 standard; protein; 510 AA.
```

```
AC AAW79740;
```

```
DT 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
```

```
XX Soybean wild-type myo-inositol 1-phosphate synthase.
```

```
KW Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KW phytic acid.
```

```
XX Glycine max; line LR13.
```

```
XX MO9845448-A1.
```

```
XX 15-OCT-1998.
```

```
XX 07-APR-1998; 98WO-US006822.
```

```
XX 08-APR-1997; 97US-00835751.
```

```
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
```

```
XX Hitz WD, Sebastian SA;
```

```
XX WPI; 1998-568353/48.
```

```
XX N-PSDB; AAV62440.
```

```
PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
PT phytic acid, etc.
```

```
XX Example 5; Page 45-47; 63pp; English.
```

```
CC This is the amino acid sequence of soybean myo-inositol 1-phosphate
CC synthase (MI 1-PS) deduced from the coding region of an isolated cDNA
CC clone (see AAV62440). MI 1-PS is involved in glucose metabolism to phytic
CC acid, raffinose and stachyose. A mutant MI 1-PS (see AAW79741) has been
CC identified in soybean line LR13, a mutagenised line of low raffinose
CC saccharide phenotype. Sequencing revealed in a single base change mutation
CC in the LR13 gene sequence that resulted in a K396N substitution in the
```

```
CC mutant protein. The mutation results in a seed phenotype of very low
CC raffinose saccharide sugars, very high sucrose and low phytic acid. The
CC mutated nucleic acid is used to alter the raffinose saccharide, sucrose,
CC phytic acid and inorganic phosphate content of soybean seeds, leading to
CC useful soybean products, e.g. a seed phytic acid content of less than 17
CC ug/g, a seed content of raffinose and stachyose combined of less than 10
CC 14.5 ug/g, and a seed sucrose content greater than 200 ug/g. (updated on
CC 17-Oct-2003 to standardise OS field)
```

```
XX Sequence 510 AA;
```

```
Query Match 98.6%; Score 2593; DB 2; Length 510;
```

```
Best Local Similarity 98.4%; Pred. No. 1,5e-212;
```

```
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 MFENFKVESPNVYKTEIETISQVYNYETTELVEHNRNGTQOMIVKPSVNYQKNTHTVP 60
DB 1 MFENFKVECPNVYKTEIETISQVYNYETTELVEHNRNGTQOMIVKPSVKKEFTNIHVP 60
QY 61 KLGYMLVGWGGNNGSTLTGGVIANREDISWATKDKIOANYFGSLTQASAIRVGSFGQEE 120
DB 61 KLGYMLVGWGGNNGSTLTGGVIANREGISWATKDKIOANYFGSLTQASAIRVGSFGQEE 120
QY 121 IYAPFKSLIPVNPDDIVFGGMDISNMNLADAMARAVFDIDLQKOLRPYMESWVPLPGI 180
DB 121 IYAPFKSLIPVNPDDIVFGGMDISNMNLADAMARAVFDIDLQKOLRPYMESWVPLPGI 180
QY 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAKFKEATKVDKVVVMTANTERSYSLV 240
DB 181 YDPDFIAANOBERANNVIKGTKEOVQOIIKDIAKFKEATKVDKVVVMTANTERSYSLV 240
QY 241 VGLNDTMENLLAAVDNRNEAISPSTLYAIACMVENVPPINGSPONTFVPGILDLAIANT 300
DB 241 VGLNDTMENLLAAVDNRNEAISPSTLYAIACMVENVPPINGSPONTFVPGILDLAIANT 300
QY 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGNMLAPQFRSKEISKS 360
DB 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIIVSYNHLGNDGNMLAPQFRSKEISKS 360
QY 361 NVDDDMNNSNAILKEPGEHPHVIVIKVPPYVGSKRAMDEYTSSEIFMGKNITVLHNTC 420
DB 361 NVDDDMNNSNAILKEPGEHPHVIVIKVPPYVGSKRAMDEYTSSEIFMGKNITVLHNTC 420
QY 421 EDSLILAAPIIIDLVLALSTRIOFKANEKGKFSFHPVATILSYLTKAPLVPPGTPVNV 480
DB 421 EDSLILAAPIIIDLVLALSTRIOFKANEKGKFSFHPVATILSYLTKAPLVPPGTPVNV 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 8

```
ADQ14499
ID ADQ14499 standard; protein; 510 AA.
```

```
AC ADQ14499;
```

```
DT 23-SEP-2004 (first entry)
```

```
XX Mutant soybean myo-inositol 1-phosphate synthase polypeptide #2.
```

```
KW Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
```

```
KW mutant; mtein.
```

```
XX Glycine max.
```

```
XX Synthetic.
```

```
XX US2004128713-A1.
```

```
XX 01-JUL-2004.
```

21-NOV-2003; 2003US-00718952.
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 PR 26-APR-1999; 99US-00299315.
 PR 11-MAR-2002; 2002US-00025003.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PA (GRAC/) GRACE D J.
 PA (STRE/) STREIT L G.
 XX
 PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
 XX
 DR WPI, 2004-533135/51.
 DR N-PSDB; ADQ14498.
 XX
 PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytic acid and increased sucrose, leading to valuable and useful soybean
 PT products.
 XX
 PS Example 8; SEQ ID NO 10; 48bp; English.

The invention relates to a nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate synthase having a decreased capacity for the synthesis of myo-inositol 1-phosphate. The invention also relates to a chimeric gene operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant comprising the chimeric gene, a method of making the soybean plant, a seed of the soybean plant, a soy protein product derived from the processing of soybean seeds, a method of making or producing a soy protein product and a method of using a soybean plant homozygous for at least one gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate. The nucleic acid fragment and methods are useful for producing plants with decreased raffinose, stachyose and phytic acid content and increased sucrose and inorganic phosphate content, leading to valuable and useful soybean products. This sequence represents a mutant soybean myo-inositol 1-phosphate synthase polypeptide of the invention.

Sequence 510 AA;

Query Match 98.6%; Score 2593; DB 8; Length 510;

Best Local Similarity 98.4%; Pred. No. 1.5e-212;

Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MFIENFKVESPVKYKTEREIQSVNYETTELVEHNRNGTYQIVPKSVNYQFKNTHTVP 60
 DB 1 MFIENFKVCEPNVKTETEIQSVNYETTELVEHNRNGTYQIVPKSVNYQFKNTHTVP 60
 QY 61 KLGWLVVGCGNGNSTLTGCVIANREDISWATKDKIQOANYGSLTQASAIWVSFGQGE 120
 DB 61 KLGWLVVGCGNGNSTLTGCVIANREDISWATKDKIQOANYGSLTQASAIWVSFGQGE 120
 QY 121 IYAPFKSLIPVNPDPDIFVCGWDISNMMLADAMAKAVPDDLOKQLRPYMESWVPLPGI 180
 DB 121 IYAPFKSLIPVNPDPDIFVCGWDISNMMLADAMAKAVPDDLOKQLRPYMESWVPLPGI 180
 QY 181 YDPDPIAANOEERANNVIKGTQOEVOQIIKQIKAFKATKVDKVVVMTANTERYSNLV 240
 DB 181 YDPDPIAANOEERANNVIKGTQOEVOQIIKQIKAFKATKVDKVVVMTANTERYSNLV 240
 QY 241 VGLNPTMENLLAAVRNEAEISPTLYAIACMEVNPFGNSPQTPVPGILDIAIART 300
 DB 241 VGLNPTMENLLAAVRNEAEISPTLYAIACMEVNPFGNSPQTPVPGILDIAIART 300
 QY 301 LIGGDDFKSGQTKMSVLDPLVGAIGKPTSIYSYNNHGGNDGMNLSAPQTFRSKEISKS 360
 DB 301 LIGGDDFKSGQTKMSVLDPLVGAIGKPTSIYSYNNHGGNDGMNLSAPQTFRSKEISKS 360

QY 361 NVVDDMVNSNAIIVRGEHPHVVIKYVPYVGDSEKRAMDEXTSEIFMGSKNTIVLHNTC 420
 DB 361 NVVDDMVNSNAIIVRGEHPHVVIKYVPYVGDSEKRAMDEXTSEIFMGSKNTIVLHNTC 420
 QY 421 EDSLAAPIIIDLVLLAEISTRIQFAENEGKFSFHPVATILSYLTRAFLVPPGTPVYN 480
 DB 421 EDSLAAPIIIDLVLLAEISTRIQFAENEGKFSFHPVATILSYLTRAFLVPPGTPVYN 480
 QY 481 ALSKORAMLEINIRACVGLAPBNMILRYK 510
 DB 481 ALSKORAMLEINIRACVGLAPBNMILRYK 510

RESULT 9

ADQ14491
 ID ADQ14491 standard; protein; 510 AA.

ADQ14491;

23-SEP-2004 (first entry)

DE Wild type soybean myo-inositol 1-phosphate synthase polypeptide #1.

KW Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;

KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme.

OS Glycine max.

PN US2004128713-A1.

PD 01-JUL-2004.

XX 21-NOV-2003; 2003US-00718952.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX 26-APR-1999; 99US-00299315.

XX 11-MAR-2002; 2002US-00025003.

PA (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.

PA (STRE/) STREIT L G.

XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;

PI WPI, 2004-533135/51.

XX N-PSDB; ADQ14490.

XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,

XX useful for producing plants with decreased raffinose, stachyose, and

XX phytic acid and increased sucrose, leading to valuable and useful soybean

XX products.

XX Claim 3; SEQ ID NO 2; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-

XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate

XX synthase having a decreased capacity for the synthesis of myo-inositol 1-

XX phosphate. The invention also relates to a chimeric gene operably linked

XX to suitable regulatory sequences, where expression of the chimeric gene

XX results in a decrease in expression of an endogenous or native gene

CC inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 510 AA;

Query Match 98.6%; Score 2593; DB 8; Length 510;

Best Local Similarity 98.4%; Pred. No. 1.5e-212; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MFENFKVESPNVKKYETETELISVYNYETTELVHNRNGTYOMIVKPSVNYQFTNTHP 60
DB 1 MFENFKVECPNVKKTETELISVYNYETTELVHNRNGTYOMIVKPSVKKEFTNTHP 60
QY 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVSPQEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDKIOQANFGSLTQASAIRVSPQEE 120
QY 121 IYAPFKSLPWNPDIVFGGMDISNNMLADAMARAKVFDIDLOKQLRPYMESWPLPGI 180
DB 121 IYAPFKSLPWNPDIVFGGMDISNNMLADAMARAKVFDIDLOKQLRPYMESWPLPGI 180
QY 181 YDPDFIANOEERANNVIKGTKEQVOQI IKDIAFKKATKVDKVVLMNTANTERYSLV 240
DB 181 YDPDFIANOEERANNVIKGTKEQVOQI IKDIAFKKATKVDKVVLMNTANTERYSLV 240
QY 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMVNVPFINSPOFTVPGIDIAIARNT 300
DB 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMVNVPFINSPOFTVPGIDIAIARNT 300
QY 301 LIGGDDFKSGQTKKKSIVLDFVGAGIKPTISIVSYNHLGNDGNMLSAPQFRSKETSKS 360
DB 301 LIGGDDFKSGQTKKKSIVLDFVGAGIKPTISIVSYNHLGNDGNMLSAPQFRSKETSKS 360
QY 361 NVVDMDMNSNAIIVRGEHPHHVVIKYVPVGSKRAMDEYTSIEFMGKNTIVLHNTC 420
DB 361 NVVDMDMNSNAIIVRGEHPHHVVIKYVPVGSKRAMDEYTSIEFMGKNTIVLHNTC 420
QY 421 EDSLILAPIIIDLVILAEISTRIOFKANEKGFSFHPVATITLSYLTAAPIVPGTPVYN 480
DB 421 EDSLILAPIIIDLVILAEISTRIOFKANEKGFSFHPVATITLSYLTAAPIVPGTPVYN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 10
ID ADS82000 standard; protein, 510 AA.
XX ADS82000;
AC ADS82000;
XX
DB 18-NOV-2004 (first entry)
XX
DE soybean myo-inositol 1-phosphate synthase wild-type 3.
XX
KM Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; line 29004JP01.
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
XX
PR 07-APR-1998; 98WO-US006822.
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
FL Hitz WD, Sebastian SA;

```

XX WPI; 2004-639957/62.
DR N-PSDB; ADS81999.

XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

XX Example 8; SEQ ID NO 10; 34pp; English.

The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a wild-type myo-inositol 1-phosphate synthase.

XX Sequence 510 AA;

Query Match 98.6%; Score 2593; DB 8; Length 510;

Best Local Similarity 98.4%; Pred. No. 1.5e-212; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MFENFKVESPNVKKYETETELISVYNYETTELVHNRNGTYOMIVKPSVNYQFTNTHP 60
DB 1 MFENFKVECPNVKKTETELISVYNYETTELVHNRNGTYOMIVKPSVKKEFTNTHP 60
QY 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIOQANFGSLTQASAIRVSPQEE 120
DB 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDKIOQANFGSLTQASAIRVSPQEE 120
QY 121 IYAPFKSLPWNPDIVFGGMDISNNMLADAMARAKVFDIDLOKQLRPYMESWPLPGI 180
DB 121 IYAPFKSLPWNPDIVFGGMDISNNMLADAMARAKVFDIDLOKQLRPYMESWPLPGI 180
QY 181 YDPDFIANOEERANNVIKGTKEQVOQI IKDIAFKKATKVDKVVLMNTANTERYSLV 240
DB 181 YDPDFIANOEERANNVIKGTKEQVOQI IKDIAFKKATKVDKVVLMNTANTERYSLV 240
QY 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMVNVPFINSPOFTVPGIDIAIARNT 300
DB 241 VGLNDTMENLLAAVDRNEAISPSTLYAIACMVNVPFINSPOFTVPGIDIAIARNT 300
QY 301 LIGGDDFKSGQTKKKSIVLDFVGAGIKPTISIVSYNHLGNDGNMLSAPQFRSKETSKS 360
DB 301 LIGGDDFKSGQTKKKSIVLDFVGAGIKPTISIVSYNHLGNDGNMLSAPQFRSKETSKS 360
QY 361 NVVDMDMNSNAIIVRGEHPHHVVIKYVPVGSKRAMDEYTSIEFMGKNTIVLHNTC 420

```


DB 361 NVVDDMVNSNALIYEPGHPDHVVYIKYVPYVGSKRAMDEYTSIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPIIIDVLVLAELSTRIOFKAENEGKFSFHPATILSLTAPLVPPTGTPVN 480
DB 421 EDSLLAAPIIIDVLVLAELSTRIEKAEKGFHSFHPATILSLTAPLVPPTGTPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
RESULT 11
ADS81994
ID ADS81994 standard; protein; 510 AA.
AC ADS81994;
XX
XX 18-NOV-2004 (first entry)
DT
XX
DE Soybean myo-inositol 1-phosphate synthase wild-type 1.
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; enzyme; phytic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; cultivar Wye.
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
PD
XX
PF 11-MAR-2002; 2002US-00025003.
PR
XX 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
DR N-PSDB; ADS81993.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
PS
XX Claim 3; SEQ ID NO 2; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for

CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence represents a wild-type myo-
CC inositol 1-phosphate synthase.
XX
SQ Sequence 510 AA:
Query Match 98.6%; Score 2593; DB 8; Length 510;
Best Local Similarity 98.4%; Pred. No..1.5e-212;
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 1 MFIENFKIESPNVKTETREIOSVYNYETTELVHNRNCTQWYVPKSVNQFTKTHVP 60
DB 1 MFIENFKIECPNVKXTEREIQSVYNYETTELVHNRNCTQWYVPKSVKVEFKTHIHP 60
QY 61 KLGVMLVGMGNGNSTLTGVIANREDISWATKDKIOQANYFGSLTQASAIRVGSFQGE 120
DB 61 KLGVMLVGMGNGNSTLTGVIANREGISWATKDKIOQANYFGSLTQASAIRVGSFQGE 120
QY 121 IYAPFKSLPMVNPDDIVFGGMDISNMNLADAMARAKVFDIDLOKLRPYMESWVPLPGI 180
DB 121 IYAPFKSLPMVNPDDIVFGGMDISNMNLADAMARAKVFDIDLOKLRPYMESWVPLPGI 180
QY 181 YDPDFIAANOEBERANNVKGTQKQOVQOIIDIAFKKATVDKRVVLTANTERYSLV 240
DB 181 YDPDFIAANOEBERANNVKGTQKQOVQOIIDIAFKKATVDKRVVLTANTERYSLV 240
QY 241 VGLNDTMENLAAVDRNEAISPSTLYAIACMENVPEPINGSPONTFPGILDLAIANT 300
DB 241 VGLNDTMENLAAVDRNEAISPSTLYAIACMENVPEPINGSPONTFPGILDLAIANT 300
QY 301 LIIGDDFKSQGTKKKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSPQTFRSKEISKS 360
DB 301 LIIGDDFKSQGTKKKSVLVDFLVGAGIKPTISIVSYNHLGNNDGMNLSPQTFRSKEISKS 360
QY 361 NVVDDMVNSNALIYEPGHPDHVVYIKYVPYVGSKRAMDEYTSIFMGKSTIVLHNTC 420
DB 361 NVVDDMVNSNALIYEPGHPDHVVYIKYVPYVGSKRAMDEYTSIFMGKSTIVLHNTC 420
QY 421 EDSLLAAPIIIDVLVLAELSTRIOFKAENEGKFSFHPATILSLTAPLVPPTGTPVN 480
DB 421 EDSLLAAPIIIDVLVLAELSTRIEKAEKGFHSFHPATILSLTAPLVPPTGTPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
RESULT 12
AAW79741
ID AAW79741 standard; protein; 510 AA.
XX
XX AAW79741;
AC
XX 17-OCT-2003 (revised)
DT
XX 02-FEB-1999 (first entry)
DT
XX
DE Soybean mutant myo-inositol 1-phosphate synthase.
XX
XX Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KM phytic acid.
XX
XX Glycine max; line LR33.
OS
XX MO9845448-A1.
PN
XX 15-OCT-1998.
PD
XX 07-APR-1998; 98WO-US006822.
PF
XX 08-APR-1997; 97US-00835751.

XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA
 XX
 XX
 PI Hitz WD, Sebastian SA;
 XX
 XX WPI; 1998-568353/48.
 DR N-PSDB; AAV62443.
 XX
 XX Soybean plants containing altered myo-inositol-1-phosphate gene - useful
 PT for generating plants with altered levels of e.g. raffinose, stachyose,
 PT phytic acid, etc.
 XX
 XX Example 5; Page 49-51; 63pp; English.

CC This is the amino acid sequence of a mutant soybean myo-inositol 1-
 CC phosphate synthase (MI 1-PS) deduced from the coding region of an
 CC isolated cDNA clone (see AAV62443). MI 1-PS is involved in glucose
 CC metabolism to phytic acid, raffinose and stachyose. The MI 1-PS was
 CC identified in soybean line LR33, a mutagenised line of low raffinose
 CC saccharide phenotype. Sequencing revealed a single base change mutation
 CC in the LR33 gene sequence that resulted in a K396N substitution in the
 CC mutant protein compared to wild-type MI 1-PS (see AAV79740). The mutation
 CC results in a seed phenotype of very low raffinose saccharide sugars, very
 CC high sucrose and low phytic acid. The mutated nucleic acid is used to
 CC alter the raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds, leading to useful soybean products,
 CC e.g. a seed phytic acid content of less than 17 ug/g, a seed content of
 CC raffinose and stachyose combined of less than 14.5 ug/g, and a seed
 CC sucrose content greater than 200 ug/g. (Updated on 17-Oct-2003 to
 CC standardise OS field)
 CC
 XX Sequence 510 AA;

Query Match 98.4%; Score 2588; DB 2; Length 510;

Best Local Similarity 98.2%; Pred. No. 4.1e-212; Mismatches 5; Indels 0; Gaps 0;

Matches 501; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 MEIENFKVESPVKYETEIEISVNYETTELVEHNRNGTYQWIVKPSVNYQFKNTHTVP 60
 DB 1 MEIENFKVCEPCVVKYTEIEISVNYETTELVEHNRNGTYQWIVKPSVNYQFKNTHTVP 60
 QY 61 KLGVMLVWVGNNSTLTGCVIANEDISWATKDIQOANTFGSLTQASAIRVSGFQGE 120
 DB 61 KLGVMLVWVGNNSTLTGCVIANEDISWATKDIQOANTFGSLTQASAIRVSGFQGE 120
 QY 61 KLGVMLVWVGNNSTLTGCVIANEDISWATKDIQOANTFGSLTQASAIRVSGFQGE 120
 DB 61 KLGVMLVWVGNNSTLTGCVIANEDISWATKDIQOANTFGSLTQASAIRVSGFQGE 120
 QY 121 IYAPKSLPWNVPDIFVGGWDISNMNLADMAKAFDIDLOQLRPVMSWVPLGI 180
 DB 121 IYAPKSLPWNVPDIFVGGWDISNMNLADMAKAFDIDLOQLRPVMSWVPLGI 180
 QY 121 IYAPKSLPWNVPDIFVGGWDISNMNLADMAKAFDIDLOQLRPVMSWVPLGI 180
 DB 121 IYAPKSLPWNVPDIFVGGWDISNMNLADMAKAFDIDLOQLRPVMSWVPLGI 180
 QY 181 YDPDFIAANOEBERANNVIGKTQEOVOQIIKIDIKAFKEATKVDKVVMTANTERYSNLV 240
 DB 181 YDPDFIAANOEBERANNVIGKTQEOVOQIIKIDIKAFKEATKVDKVVMTANTERYSNLV 240
 QY 241 VGLNDTMENLAAVDRNEAISPSTLYAIACMVENVPINGSPOPTFPGGIDILAIANT 300
 DB 241 VGLNDTMENLAAVDRNEAISPSTLYAIACMVENVPINGSPOPTFPGGIDILAIANT 300
 QY 241 VGLNDTMENLAAVDRNEAISPSTLYAIACMVENVPINGSPOPTFPGGIDILAIANT 300
 DB 241 VGLNDTMENLAAVDRNEAISPSTLYAIACMVENVPINGSPOPTFPGGIDILAIANT 300
 QY 301 LIIGDGFSSGOTKMSVIVDFVAGIKPTSVSVNHHGNDGNMLSAPOFRSKETSKS 360
 DB 301 LIIGDGFSSGOTKMSVIVDFVAGIKPTSVSVNHHGNDGNMLSAPOFRSKETSKS 360
 QY 301 LIIGDGFSSGOTKMSVIVDFVAGIKPTSVSVNHHGNDGNMLSAPOFRSKETSKS 360
 DB 301 LIIGDGFSSGOTKMSVIVDFVAGIKPTSVSVNHHGNDGNMLSAPOFRSKETSKS 360
 QY 361 NVVDVMVSNAILYEPGEHPDHVVVVKVVPYVDSKRAMDETSRIFMGKNTVILANTC 420
 DB 361 NVVDVMVSNAILYEPGEHPDHVVVVKVVPYVDSKRAMDETSRIFMGKNTVILANTC 420
 QY 421 EDLSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHVATITLSYLTAAPLVPCTPVVN 480
 DB 421 EDLSLAAPIIIDLVLLAELSTRIOFKANEKGKFSFHVATITLSYLTAAPLVPCTPVVN 480
 QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
 DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 13
 ADQ14495
 ID ADQ14495 standard; protein; 510 AA.

XX AC ADQ14495;

XX DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase polypeptide #1.

XX Soybean; myo-inositol 1-phosphate synthase; myo-inositol 1-phosphate;
 KW raffinose; stachyose; phytic acid; sucrose; inorganic phosphate; enzyme;
 KW mutant; mutain.

XX OS Glycine max.
 XX Synthetic.

XX Key Location/Qualifiers
 FT Misc-difference 396
 FT /note= "Wild type Lys substituted by Asn"

XX US2004128713-A1.

XX PD 01-JUL-2004.

XX PF 21-NOV-2003; 2003US-00718952.

XX PR 08-APR-1997; 97US-00835751.

XX PR 07-APR-1998; 98WO-US006822.

XX PR 26-APR-1999; 99US-0029315.

XX PR 11-MAR-2002; 2002US-00025003.

XX PA (HITZ/) HITZ W D.

XX PA (SRBA/) SRBASTIAN S A.

XX PA (GRAC/) GRACE D J.

XX PA (STRE/) STREIT L G.

XX PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;

XX PS WPI; 2004-533135/51.

XX DR N-PSDB; ADQ14494.

XX PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytic acid and increased sucrose, leading to valuable and useful soybean
 PT products.

XX PR Claim 9; SEQ ID NO 6; 48pp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
 CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
 CC phosphate. The invention also relates to a chimeric gene operably linked
 CC to suitable regulatory sequences, where expression of the chimeric gene
 CC results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
 CC comprising the chimeric gene, a method of making the soybean plant, a
 CC seed of the soybean plant, a soy protein product derived from the
 CC processing of soybean seeds, a method of making or producing a soy
 CC protein product and a method of using a soybean plant homozygous for at
 CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 CC nucleic acid fragment and methods are useful for producing plants with
 CC decreased raffinose, stachyose and phytic acid content and increased
 CC sucrose and inorganic phosphate content, leading to valuable and useful
 CC soybean products. This sequence represents a mutant soybean myo-inositol
 CC 1-phosphate synthase polypeptide of the invention.

XX Sequence 510 AA;

Query Match 98.4%; Score 2588; DB 8; Length 510;
 Best Local Similarity 98.2%; Pred. No. 4.1e-212;

Matches	501;	Conservative	4;	Mismatches	5;	Indels	0;	Gaps	0;
QY	1	MEIENFKVESPVVKYTEREIQSVNYETTELVEHNRNGTYOMIVKPKSVNYQFKNTNHP	60						
DB	1	MEIENFKVESPVVKYTEREIQSVNYETTELVEHNRNGTYOMIVKPKSVNYQFKNTNHP	60						
QY	61	KLGVNLVGMGNGNSTLTGVIANREDISWATKDKIQOANFGSLTQASAIRVGSFOGEE	120						
DB	61	KLGVNLVGMGNGNSTLTGVIANREDISWATKDKIQOANFGSLTQASAIRVGSFOGEE	120						
QY	121	IYAPFKSLPVMNPPDDIVFGGWDISNNMLADAMARAKYFDIDLOKOLRPYMESMPLPGI	180						
DB	121	IYAPFKSLPVMNPPDDIVFGGWDISNNMLADAMARAKYFDIDLOKOLRPYMESMPLPGI	180						
QY	181	YDPDFIAANOEBRANNVIGTQEOVOQI IKDIAKFAKATVDKVVVMTANTERYSLV	240						
DB	181	YDPDFIAANOEBRANNVIGTQEOVOQI IKDIAKFAKATVDKVVVMTANTERYSLV	240						
QY	241	VGNDTMENTLAAVDRNEAISPSTLYAIACVMENVPFINSPONTFVPGIDLAIAANT	300						
DB	241	VGNDTMENTLAAVDRNEAISPSTLYAIACVMENVPFINSPONTFVPGIDLAIAANT	300						
QY	301	LIGGDDFKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNDGNMLSAPQFRSKISKS	360						
DB	301	LIGGDDFKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNDGNMLSAPQFRSKISKS	360						
QY	361	NVDDMVNSNALIYEPGHPDHVVIVKVPYVGSKRAMEYTSIEIFMGKNTIVLANTC	420						
DB	361	NVDDMVNSNALIYEPGHPDHVVIVKVPYVGSKRAMEYTSIEIFMGKNTIVLANTC	420						
QY	421	EDSLIAAPILIDVLAELSTRIOFKARENGKFSFHPATILSLVLTAPLPVPGTPPVN	480						
DB	421	EDSLIAAPILIDVLAELSTRIOFKARENGKFSFHPATILSLVLTAPLPVPGTPPVN	480						
QY	481	ALSKORAMLENIMRACVGLAPENNMILEYK	510						
DB	481	ALSKORAMLENIMRACVGLAPENNMILEYK	510						

XX	WPI, 2004-639957/62.
DR	N-PsDB; ADS81997.
PT	Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
PS	Claim 9, SEQ ID NO 6, 34pp; English.
CC	The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/9, a seed content of raffinose plus stachyose of less than 14.5 micromol/9, and a seed sucrose content of greater than 200 micromol/9, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence represents a mutant myo-inositol 1-phosphate synthase.
SO	Sequence 510 AA:
QY	1 MEIENFKVESPVVKYTEREIQSVNYETTELVEHNRNGTYOMIVKPKSVNYQFKNTNHP
DB	1 MEIENFKVESPVVKYTEREIQSVNYETTELVEHNRNGTYOMIVKPKSVNYQFKNTNHP
QY	61 KLGVNLVGMGNGNSTLTGVIANREDISWATKDKIQOANFGSLTQASAIRVGSFOGEE
DB	61 KLGVNLVGMGNGNSTLTGVIANREDISWATKDKIQOANFGSLTQASAIRVGSFOGEE
QY	121 IYAPFKSLPVMNPPDDIVFGGWDISNNMLADAMARAKYFDIDLOKOLRPYMESMPLPGI
DB	121 IYAPFKSLPVMNPPDDIVFGGWDISNNMLADAMARAKYFDIDLOKOLRPYMESMPLPGI
QY	181 YDPDFIAANOEBRANNVIGTQEOVOQI IKDIAKFAKATVDKVVVMTANTERYSLV
DB	181 YDPDFIAANOEBRANNVIGTQEOVOQI IKDIAKFAKATVDKVVVMTANTERYSLV
QY	241 VGNDTMENTLAAVDRNEAISPSTLYAIACVMENVPFINSPONTFVPGIDLAIAANT
DB	241 VGNDTMENTLAAVDRNEAISPSTLYAIACVMENVPFINSPONTFVPGIDLAIAANT
QY	301 LIGGDDFKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNDGNMLSAPQFRSKISKS
DB	301 LIGGDDFKSGQTKKSVLVDFLVGAGIKPTISIVSYNHLGNDGNMLSAPQFRSKISKS
QY	361 NVDDMVNSNALIYEPGHPDHVVIVKVPYVGSKRAMEYTSIEIFMGKNTIVLANTC

Query Match 98.4%; Score 2588; DB 8; Length 510;
 Best Local Similarity 98.2%; Pred. No. 4.1e-212;
 Matches 501; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 361 NVVDDMNSNALIYEPGHPDHVVVIKYVPYVGDNSRAMDEYTSIEFMGKSTIVLHNTC 420
QY 421 EDSLAAPIIIDLVLAELSTRIOKANEKGKFSFHPVATILSYLTKAPLVPPTPVVN 480
Db 421 EDSLAAPIIIDLVLAELSTRIOKANEKGKFSFHPVATILSYLTKAPLVPPTPVVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 15

AAV24477
ID AAV24477 standard; protein; 536 AA.

AAV24477;
AC
XX
XX

DT 24-SEP-1999 (first entry)

XX Nicotiana paniculata INPS protein.

XX Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;
KW water stress; resistance.

XX Nicotiana paniculata.

XX Key Location/Qualifiers

FT Misc-difference 511 /label= unknown
FT /note= "encoded by the stop codon TGA"

XX JP1187879-A.

XX 13-JUL-1999.

XX 26-DEC-1997; 97JP-00359773.

XX 26-DEC-1997; 97JP-00359773.

XX (NISH) JAPAN TOBACCO INC.

XX WPI, 1999-451546/38.

XX N-PSDB; AAX90402.

XX New INPS gene derived from Nicotiana genus plant - useful for conferring
PT resistance to water stress to plants.

XX Claim 2; Page 6-8; 8bp; Japanese.

CC The present sequence is the Nicotiana paniculata inositol monophosphate
CC synthase (INPS), designated NpINPS1. INPS can be used to confer water
CC stress resistance to a plant

XX Sequence 536 AA;

Query Match 93.9%; Score 2470; DB 2; Length 536;
Best Local Similarity 92.4%; Pred. No. 5,4e-202;

Matches 471; Conservative 21; Mismatches 18; Indels 0; Gaps 0;

QY 1 MFIENFKVESPVVKKTETETIEIQSYVNYETTELVDHENRNGTYQWIVPKSVNTQFKTNTVVP 60
Db 1 MFIENFKVESPVVKKTETETIEIQSYVNYETTELVDHENRNGTYQWIVPKSVNTQFKTNTVVP 60
QY 61 KLGVVLVGMGNGNSTLTGTVANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
Db 61 KLGVVLVGMGNGNSTLTGTVANREDISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKVFDIDLQKQLRPYMESWVPLPGI 180
Db 121 IYAPFKSLIPVNPDDIVFGGWDISNNMLADAMARAKVFDIDLQKQLRPYMESWVPLPGI 180
QY 181 YDDPFIANQGBRANNVIKGTKEQIDQIILKDIREFEKKKVKVVLWTANTERYSNV 240
Db 181 YDDPFIANQGBRANNVIKGTKEQIDQIILKDIREFEKKKVKVVLWTANTERYSNV 240

Db 181 YDDPFIANQGBRANNVIKGTKEQIDQIILKDIREFEKKKVKVVLWTANTERYSNV 240
QY 241 VGLNDTMENTLAAVDRNEAEISPTLYAIAQWENVPFINGSPONTFVPGILDLAIAANT 300
Db 241 VGLNDTMENTLAAVDRNEAEISPTLYAIAQWENVPFINGSPONTFVPGILDLAIAKNT 300
QY 301 LIGGDDFKSGQTKKKSIVLVFVGAGIKPTISIVSYNHLGNNDGNLAPQTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKKKSIVLVFVGAGIKPTISIVSYNHLGNNDGNLAPQTFRSKEISKS 360
QY 361 NVVDDMNSNALIYEPGHPDHVVVIKYVPYVGDNSRAMDEYTSIEFMGKNTIVLHNTC 420
Db 361 NVVDDMNSNALIYEPGHPDHVVVIKYVPYVGDNSRAMDEYTSIEFMGKNTIVLHNTC 420
QY 421 EDSLAAPIIIDLVLAELSTRIOKANEKGKFSFHPVATILSYLTKAPLVPPTPVVN 480
Db 421 EDSLAAPIIIDLVLAELSTRIOKANEKGKFSFHPVATILSYLTKAPLVPPTPVVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

Search completed: June 7, 2005, 16:43:37
Job time : 93.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 03:03:45 ; Search time 952.946 Seconds
(without alignments)
9972.402 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533

Sequence: 1 acgtccatcagagaattctaa.....tgattctcaggracaagtga 1533

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1533	100.0	1533	14	US-10-025-003-11
2	1533	100.0	1533	19	US-10-718-952-11
3	1531.4	99.9	1533	14	US-10-025-003-11
4	1531.4	99.9	1533	14	US-10-025-003-11
5	1531.4	99.9	1533	19	US-10-718-952-11
6	1531.4	99.9	1533	19	US-10-718-952-11
7	1464.2	95.5	1533	14	US-10-025-003-9

8	1464.2	95.5	1533	19	US-10-718-952-9	Sequence 9, Appl1
9	1464.2	95.5	1760	14	US-10-025-003-1	Sequence 1, Appl1
10	1464.2	95.5	1760	14	US-10-718-952-1	Sequence 1, Appl1
11	1462.6	95.4	1533	14	US-10-025-003-5	Sequence 5, Appl1
12	1462.6	95.4	1533	19	US-10-718-952-5	Sequence 5, Appl1
13	1456.4	95.0	1989	18	US-10-424-599-70167	Sequence 70167, A
14	1261	82.3	2018	18	US-10-424-599-12022	Sequence 12022, A
15	1176.4	76.7	2582	18	US-10-424-599-12021	Sequence 12021, A
16	1000.2	65.2	1924	19	US-10-767-701-14714	Sequence 14714, A
17	998.6	65.1	1533	9	US-09-938-842A-1438	Sequence 1438, Ap
18	998.6	65.1	1533	11	US-09-938-842A-1438	Sequence 1438, Ap
19	995.4	64.9	1941	20	US-10-739-930-2091	Sequence 2091, Ap
20	990.6	64.6	2284	20	US-10-425-115-47189	Sequence 47189, A
21	989	64.5	2156	18	US-10-425-115-122151	Sequence 25858, A
22	987.4	64.4	2444	20	US-10-425-115-122151	Sequence 122151, A
23	985.8	64.3	1988	18	US-10-425-114-7081	Sequence 7081, Ap
24	984.2	64.2	1879	19	US-10-425-114-5704	Sequence 5704, Ap
25	984.2	64.2	1914	18	US-10-425-114-5612	Sequence 5612, Ap
26	984.2	64.2	1959	9	US-09-727-628-1	Sequence 1, Appl1
27	982.6	64.1	1971	18	US-10-425-114-26983	Sequence 26983, A
28	982.6	64.1	2233	20	US-10-425-115-47187	Sequence 47187, A
29	981	64.0	2309	20	US-10-425-115-47194	Sequence 47194, A
30	981	64.0	2576	19	US-10-437-963-70769	Sequence 70769, A
31	979.4	63.9	1912	18	US-10-425-114-8483	Sequence 8483, Ap
32	979.4	63.9	1931	9	US-09-921-232-10	Sequence 10, Appl1
33	979.4	63.9	1931	9	US-09-921-330-10	Sequence 10, Appl1
34	979.4	63.9	1931	9	US-09-921-329-10	Sequence 10, Appl1
35	979.4	63.9	1997	20	US-10-425-115-47190	Sequence 47190, A
36	970	63.3	2582	18	US-10-424-599-12021	Sequence 12021, A
37	941.6	61.4	1759	19	US-10-363-889-205	Sequence 205, App
38	926.6	60.4	1858	18	US-10-425-114-14009	Sequence 14009, A
39	926.6	60.4	2050	18	US-10-425-114-5864	Sequence 5864, Ap
40	925	60.3	1853	18	US-10-425-114-16524	Sequence 16524, Ap
41	925	60.3	2130	20	US-10-425-115-83553	Sequence 83553, A
42	910.8	59.4	1673	20	US-10-425-115-47180	Sequence 47180, A
43	901	58.8	2196	19	US-10-437-963-85285	Sequence 85285, A
44	891.6	58.2	1789	20	US-10-425-115-47182	Sequence 47182, A
45	665	43.4	1240	18	US-10-425-114-5619	Sequence 5619, Ap

ALIGNMENTS

RESULT 1
US-10-025-003-11
; Sequence 11, Application US/10025003
; Publication No. US2003007465A1
GENERAL INFORMATION:
APPLICANT: Hitz, William
APPLICANT: Sebastian, Scott
APPLICANT: Grace, John
APPLICANT: Streit, Leon
TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
FILE REFERENCE: BB-1077-C
CURRENT APPLICATION NUMBER: US/10/025, 003
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 08/835, 751
PRIOR FILING DATE: APRIL 8, 1997
PRIOR APPLICATION NUMBER: PCT/US98/06822
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Microsoft Office 97
SEQ ID NO 11
LENGTH: 1533
TYPE: DNA
ORGANISM: Glycine max
US-10-025-003-11
Query Match 100.0%; Score 1533; DB 14; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 ATGTCATCGAGAAATTTTAAGGTAGAGAGTCTTAATGTGAATGACACCGAGACTGAGATT 60
DB 1 ATGTTCAATCGAGAAATTTTAAGGTAGAGAGTCTTAATGTGAATGACACCGAGACTGAGATT 60
OY 61 CAGTCGCTGTACAACTACGAAACCAACCGAATTGTTTCAGAGAACAGAAATGGCACTAT 120
DB 61 CAGTCGCTGTACAACTACGAAACCAACCGAATTGTTTCAGAGAACAGAAATGGCACTAT 120
OY 121 CAGTGGATTTGTCAAAACCCAAATCCGTCACATCAATTTAAACCAACACCCCATGTGTCA 180
DB 121 CAGTGGATTTGTCAAAACCCAAATCCGTCACATCAATTTAAACCAACACCCCATGTGTCA 180
OY 181 AAATGGGGGTGATGCTTGTGGGTTGGGGTGGAAAACAAGGCTCTACCTTCACCGGTGCT 240
DB 181 AAATGGGGGTGATGCTTGTGGGTTGGGGTGGAAAACAAGGCTCTACCTTCACCGGTGCT 240
OY 241 GTTATTTGCTAACAGAGAGACATTTTCATGGGCTACAAAGAGACAAATTCACAAAGCCAT 300
DB 241 GTTATTTGCTAACAGAGAGACATTTTCATGGGCTACAAAGAGACAAATTCACAAAGCCAT 300
OY 301 TACTTTGGCTCCCTCAACCCAGCTTCAGTATTCAGATTGATCCTTCCAGGGAGAGAA 360
DB 301 TACTTTGGCTCCCTCAACCCAGCTTCAGTATTCAGATTGATCCTTCCAGGGAGAGAA 360
OY 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCATGGTTAATCTCGACACATTTGTTGGG 420
DB 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCATGGTTAATCTCGACACATTTGTTGGG 420
OY 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCAGTGGCAAGGAGGTTTGAAC 480
DB 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCCAGTGGCAAGGAGGTTTGAAC 480
OY 481 ATCGATTTTCAGAGAGAGATTTGAGGCTTTACATGGAATTCATGGTTCACCTCCCGGAATC 540
DB 481 ATCGATTTTCAGAGAGAGATTTGAGGCTTTACATGGAATTCATGGTTCACCTCCCGGAATC 540
OY 541 TACGACCCCGGATTTCTTGTCTGCAACCAAGAGACGTCGCAACCAACGTTAAGAGGCG 600
DB 541 TACGACCCCGGATTTCTTGTCTGCAACCAAGAGACGTCGCAACCAACGTTAAGAGGCG 600
OY 601 ACAAAAGCAAGAGCAATTCAGCAAAATCATCAAAAGATCAAGGCGTTTAAAGAGCCACC 660
DB 601 ACAAAAGCAAGAGCAATTCAGCAAAATCATCAAAAGATCAAGGCGTTTAAAGAGCCACC 660
OY 661 AAAGTGAACAAGGTGTTGCTCCGTGAGACGTCGCAACCAAGAGGATATGCAATTTGGTT 720
DB 661 AAAGTGAACAAGGTGTTGCTCCGTGAGACGTCGCAACCAAGAGGATATGCAATTTGGTT 720
OY 721 GTAGGCTTAATGACACATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTAATGACACATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAG 780
OY 781 ATTTCTCTTCCACCTTGTATGTCATGCTGTGTGATGGAATAATGTTCTTTCATTAAT 840
DB 781 ATTTCTCTTCCACCTTGTATGTCATGCTGTGTGATGGAATAATGTTCTTTCATTAAT 840
OY 841 GGAAGCCCTCAGAACTTTTGTACCAAGGCTGATGATCTTCCATCGGAGAAACACT 900
DB 841 GGAAGCCCTCAGAACTTTTGTACCAAGGCTGATGATCTTCCATCGGAGAAACACT 900
OY 901 TTGATTTGTGAGATGACTTCAAGAGTGTCAAGCCAAATGAAATCTGTGTTGGTTGAT 960
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTCAAGCCAAATGAAATCTGTGTTGGTTGAT 960
OY 961 TTTCTTGTGGGGCTGTATCAAGCCAACTGTATAGTTAAGTTAACAACATCTGGGAAAC 1020
DB 961 TTTCTTGTGGGGCTGTATCAAGCCAACTGTATAGTTAAGTTAACAACATCTGGGAAAC 1020
OY 1021 AATGATGATGATGATCTCTGGGCTCCAAACCTTCCGCTCCAAAGAAATCTCAAGAGC 1080
DB 1021 AATGATGATGATGATCTCTGGGCTCCAAACCTTCCGCTCCAAAGAAATCTCAAGAGC 1080
OY 1081 AAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 AAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
```

```
DB 1081 AACGTTGTGACGATATGCTCAACAGCAATGCAATCTCTATGAGCTGTGAAATCCC 1140
OY 1141 GACCAATGTTGTTAATTAAGATGTGCTTACGATGGGAGATGACAGAGCCATGAT 1200
DB 1141 GACCAATGTTGTTAATTAAGATGTGCTTACGATGGGAGATGACAGAGCCATGAT 1200
OY 1201 GAGTACACTTCAAGATATTCATGGGTGAAAGAACACCATTTGTTTTCACAAACATGT 1260
DB 1201 GAGTACACTTCAAGATATTCATGGGTGAAAGAACACCATTTGTTTTCACAAACATGT 1260
OY 1261 GAGATTCCTTTTATGCTGCTCTATTTATCTTGAATCTTGTCCTTCTGAGCTGAGC 1320
DB 1261 GAGATTCCTTTTATGCTGCTCTATTTATCTTGAATCTTGTCCTTCTGAGCTGAGC 1320
OY 1321 ACTAATATCCAGTTTAAAGCTGAAATAGGAGAAATTCACATCTCAATCCACAGTTGCT 1380
DB 1321 ACTAATATCCAGTTTAAAGCTGAAATAGGAGAAATTCACATCTCAATCCACAGTTGCT 1380
OY 1381 ACCATTCAGCTATCTGACCAAGGCTCCTGTGTTCCACCGGGTACACCACTGTGAT 1440
DB 1381 ACCATTCAGCTATCTGACCAAGGCTCCTGTGTTCCACCGGGTACACCACTGTGAT 1440
OY 1441 GCATTGTCAAGACGCTGCAATGCTGAAAACATTAATGAGGCTTGTGTTGGCTC 1500
DB 1441 GCATTGTCAAGACGCTGCAATGCTGAAAACATTAATGAGGCTTGTGTTGGCTC 1500
OY 1501 CCAGGAATTAACATGATCTCGAGTACAAAGTGA 1533
DB 1501 CCAGGAATTAACATGATCTCGAGTACAAAGTGA 1533

RESULT 2
US-10-718-952-11
; Sequence 11, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-718-952-11

Query Match 100.0%; Score 1533; DB 19; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGTCATCGAGAAATTTTAAGGTAGAGAGTCTTAATGTGAATGACACCGAGACTGAGATT 60
DB 1 ATGTTCAATCGAGAAATTTTAAGGTAGAGAGTCTTAATGTGAATGACACCGAGACTGAGATT 60
OY 61 CAGTCGCTGTACAACTACGAAACCAACCGAATTGTTTCAGAGAACAGAAATGGCACTAT 120
DB 61 CAGTCGCTGTACAACTACGAAACCAACCGAATTGTTTCAGAGAACAGAAATGGCACTAT 120
OY 121 CAGTGGATTTGTCAAAACCCAAATCCGTCACATCAATTTAAACCAACACCCCATGTGTCA 180
DB 121 CAGTGGATTTGTCAAAACCCAAATCCGTCACATCAATTTAAACCAACACCCCATGTGTCA 180
```

QY 181 AAATTGGGGGTATGCTTGGGGTTGGGGTGAACAAGGCTCTACCCCTCACCGGTGAT 240
 DB 181 AAATTGGGGGTATGCTTGGGGTTGGGGTGAACAAGGCTCTACCCCTCACCGGTGAT 240
 QY 241 GTTATTGCTTACAGAGAGACATTTTCATGAGGCTTACAAAGACAAAGATTCAACAGCCAA 300
 DB 241 GTTATTGCTTACAGAGAGACATTTTCATGAGGCTTACAAAGACAAAGATTCAACAGCCAA 300
 QY 301 TACTTTGGCTCCCTCACCCAAAGCTCAGCTATTGAGTTGGATCTTTCACAGGAGAGAA 360
 DB 301 TACTTTGGCTCCCTCACCCAAAGCTCAGCTATTGAGTTGGATCTTTCACAGGAGAGAA 360
 QY 361 ATCTATGCCCCATTCAAGAGCTGCTTCCAAATGTTAACTCCGACCAATTGTTGGG 420
 DB 361 ATCTATGCCCCATTCAAGAGCTGCTTCCAAATGTTAACTCCGACCAATTGTTGGG 420
 QY 421 GGATGGATATACGACAACTGAACCTGCTGATGCGCATGGCCAGGCAAGGTTTGAAC 480
 DB 421 GGATGGATATACGACAACTGAACCTGCTGATGCGCATGGCCAGGCAAGGTTTGAAC 480
 QY 481 ATCGATTTCAGAGACGTTGAGGCTTTCATGGAATTCATGAGTTCCATCCCGGAAATC 540
 DB 481 ATCGATTTCAGAGACGTTGAGGCTTTCATGGAATTCATGAGTTCCATCCCGGAAATC 540
 QY 541 TACGACCCGGATTTCTTCTGCTCCAAAGAGAGAGGCTCCAAACAGTGAATTAAGGCG 600
 DB 541 TACGACCCGGATTTCTTCTGCTCCAAAGAGAGAGGCTCCAAACAGTGAATTAAGGCG 600
 QY 601 ACAAAGACAAAGACAAATTCAGCAAAATCATCAAGACATCAAGGCTTTTAAGAGAGCACC 660
 DB 601 ACAAAGACAAAGACAAATTCAGCAAAATCATCAAGACATCAAGGCTTTTAAGAGAGCACC 660
 QY 661 AAAGTGAACAAGGTGTTCTCTGCTGAGTCCCAACAGAGAGATTAAGCAATTTGGTT 720
 DB 661 AAAGTGAACAAGGTGTTCTCTGCTGAGTCCCAACAGAGAGATTAAGCAATTTGGTT 720
 QY 721 GTAGGACCTTAATGACACCAATGAGAACTCTTGGCTGCTGAGACAGAAATGAGGTGAG 780
 DB 721 GTAGGACCTTAATGACACCAATGAGAACTCTTGGCTGCTGAGACAGAAATGAGGTGAG 780
 QY 781 ATTTCTCTTCCACCTTGTATGCCATGCTGCTGATGGAATAATTTCTTCAATTAAT 840
 DB 781 ATTTCTCTTCCACCTTGTATGCCATGCTGCTGATGGAATAATTTCTTCAATTAAT 840
 QY 841 GGAAGCCCTCAGAAACCTTTGTACCAAGGCTGATTTGATCTTGCATGCGAGAACT 900
 DB 841 GGAAGCCCTCAGAAACCTTTGTACCAAGGCTGATTTGATCTTGCATGCGAGAACT 900
 QY 901 TTGATTTGGAGATGACTTCAAGAGTGTGACCAAAATGAAATCTGTGTTGGTGAAT 960
 DB 901 TTGATTTGGAGATGACTTCAAGAGTGTGACCAAAATGAAATCTGTGTTGGTGAAT 960
 QY 961 TTTCTTGGGGGGTGGTATCAAGCAACATCATATGTTAACAACCATGCGGAAAC 1020
 DB 961 TTTCTTGGGGGGTGGTATCAAGCAACATCATATGTTAACAACCATGCGGAAAC 1020
 QY 1021 AATGATGATGATATCTCGGCTCCAAACCTTCCGCTCCAAAGAAATCTTCAAGAGC 1080
 DB 1021 AATGATGATGATATCTCGGCTCCAAACCTTCCGCTCCAAAGAAATCTTCAAGAGC 1080
 QY 1081 AACGTTTGGAGATGATGATCAACAGCAATGCTCTTATGAGGCTGTGTGAACATCCC 1140
 DB 1081 AACGTTTGGAGATGATGATCAACAGCAATGCTCTTATGAGGCTGTGTGAACATCCC 1140
 QY 1141 GACCATGTTGTTATTAAGATGCTTACGTAAGGAGATGACAGAGAGCATTGAT 1200
 DB 1141 GACCATGTTGTTATTAAGATGCTTACGTAAGGAGATGACAGAGAGCATTGAT 1200
 QY 1201 GAGTACACTTCAGAGATATTCATGAGTGAAGAAACCACTTTGTTGCAACAACATGT 1260
 DB 1201 GAGTACACTTCAGAGATATTCATGAGTGAAGAAACCACTTTGTTGCAACAACATGT 1260

QY 1261 GAGATTCCTTTTATGAGTCTCTTATATCTTGAAGTGTGCTTCTGAGTGAAC 1320
 DB 1261 GAGATTCCTTTTATGAGTCTCTTATATCTTGAAGTGTGCTTCTGAGTGAAC 1320
 QY 1321 ACTAGAAATCAGTTTAAAGCTGAAAAATGAGGAAAAATTCACCTATTCACCCAGTTGCT 1380
 DB 1321 ACTAGAAATCAGTTTAAAGCTGAAAAATGAGGAAAAATTCACCTATTCACCCAGTTGCT 1380
 QY 1381 ACCATTCTCAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGCTACACCACTGTGAAT 1440
 DB 1381 ACCATTCTCAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGCTACACCACTGTGAAT 1440
 QY 1441 GCAATGTCAAGACAGGTCATGCTGAAAAATTAATGAGGCTTGTGGAATTGGGC 1500
 DB 1441 GCAATGTCAAGACAGGTCATGCTGAAAAATTAATGAGGCTTGTGGAATTGGGC 1500
 QY 1501 CCAGAGAAATACATGATTTCTCGAGTACAAAGTGA 1533
 DB 1501 CCAGAGAAATACATGATTTCTCGAGTACAAAGTGA 1533

RESULT 3
 US-10-025-003-13
 ; Sequence 13, Application US/10025003
 ; Publication No. US20030074685A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hiltz, William
 ; APPLICANT: Sebastian, Scott
 ; APPLICANT: Grace, John
 ; APPLICANT: Streif, Leon
 ; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
 ; FILE REFERENCE: BB-1077-C
 ; CURRENT APPLICATION NUMBER: US/10/025,003
 ; PRIOR FILING DATE: 2002-05-07
 ; PRIOR APPLICATION NUMBER: 08/835,751
 ; PRIOR FILING DATE: APRIL 8, 1997
 ; PRIOR APPLICATION NUMBER: PCT/US98/06822
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 13
 ; LENGTH: 1533
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; US-10-025-003-13

Query Match 99.9%; Score 1531.4; DB 14; Length 1533;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTTCATCGAGAATTTTAAGTAGAGAGTCTTAATGTAAGTACACCGAGATGAGATT 60
 DB 1 ATGTTTCATCGAGAATTTTAAGTAGAGAGTCTTAATGTAAGTACACCGAGATGAGATT 60
 QY 61 CAGTCCGTTGATCACTACGAAACCAACCGAATCTTGTTCACGAGAACAGAAATGCACTAT 120
 DB 61 CAGTCCGTTGATCACTACGAAACCAACCGAATCTTGTTCACGAGAACAGAAATGCACTAT 120
 QY 121 CAGTGGATTTGCAAAACCAAAATCCGTCAACTACCAATTTTAAACCAACCAATGTTCA 180
 DB 121 CAGTGGATTTGCAAAACCAAAATCCGTCAACTACCAATTTTAAACCAACCAATGTTCA 180
 QY 181 AAATTGGGGGTGATCTTGTGGTTGGGGTGAACAACCGCTTACCTCAACCGGTGT 240
 DB 181 AAATTGGGGGTGATCTTGTGGTTGGGGTGAACAACCGCTTACCTCAACCGGTGT 240
 QY 241 GTTATTGCTTACAGAGAGACATTTTCATGAGGCTTACAAAGACAAAGATTCAACAACCAAT 300
 DB 241 GTTATTGCTTACAGAGAGACATTTTCATGAGGCTTACAAAGACAAAGATTCAACAACCAAT 300
 QY 301 TACTTTGGCTCCCTCACCCAAAGCTCAGCTATTGAGTTGGATCTTCCAGGAGAGGAA 360

```
Db 301 TACTTGGCTCCCTCAACCCAGGCTCAGCTATTCAGATTGATCCTTCCAGGAGAGAA 360
Qy ATCTATGCCCATTCAGAGAGTCTGCTTCCATAGTGAATCTGACGACATGTGGTGGG 420
Db 361 ATCTATGCCCATTCAGAGAGTCTGCTTCCATAGTGAATCTGACGACATGTGGTGGG 420
Qy 421 GGATGGGATATCAGCAACATGAACTTGGCTGATGCGCATGGCCAGGCAAAAGGTGTTGAC 480
Db 421 GGATGGGATATCAGCAACATGAACTTGGCTGATGCGCATGGCCAGGCAAAAGGTGTTGAC 480
Qy 481 ATCGATTTCCAGAGAGAGTTCAGGCGCTTACATGGAATTCATGGTTCCACTCCCGGGAATC 540
Db 481 ATCGATTTCCAGAGAGAGTTCAGGCGCTTACATGGAATTCATGGTTCCACTCCCGGGAATC 540
Qy 541 TACGACCCCGGATTTCTTGTCTGCGCAACCAAGAGAGCGTCCCAACAACGTGATTAAGGCGC 600
Db 541 TACGACCCCGGATTTCTTGTCTGCGCAACCAAGAGAGCGTCCCAACAACGTGATTAAGGCGC 600
Qy 601 ACAAGCAAGAGAGAGTTCAGCAAAATCATCAAGACATCAAGGCGTTTAAAGAGCCACC 660
Db 601 ACAAGCAAGAGAGAGTTCAGCAAAATCATCAAGACATCAAGGCGTTTAAAGAGCCACC 660
Qy 661 AAGATGGAACAAGGTGTTGCTTCCCTGAGACTGCGCAACAGAGAGGTATAGCAATTTGGTT 720
Db 661 AAGATGGAACAAGGTGTTGCTTCCCTGAGACTGCGCAACAGAGAGGTATAGCAATTTGGTT 720
Qy 721 GTAGGCTTAATGACACCATGAGAACTCTTGGCTGCTGTGGAAGAATGAGGCTGAG 780
Db 721 GTAGGCTTAATGACACCATGAGAACTCTTGGCTGCTGTGGAAGAATGAGGCTGAG 780
Qy 781 ATTTCTCTTCCACCTTGTATGCGCATTCCTGTGATGAGAAATGTTCTTTCAATTAAT 840
Db 781 ATTTCTCTTCCACCTTGTATGCGCATTCCTGTGATGAGAAATGTTCTTTCAATTAAT 840
Qy 841 GGAAGCCCTCAGAAACCTTTTGACAGAGGCTGATTTGATCTTGGCTCCCGAGGAACT 900
Db 841 GGAAGCCCTCAGAAACCTTTTGACAGAGGCTGATTTGATCTTGGCTCCCGAGGAACT 900
Qy 901 TTGATGTTGAGAGATACCTTCAAGAGTGGTCAGAACCAATGAATCTGTGTTGGTGAAT 960
Db 901 TTGATGTTGAGAGATACCTTCAAGAGTGGTCAGAACCAATGAATCTGTGTTGGTGAAT 960
Qy 961 TTTCTTGGGGGCTGATCAAGCAACATCTATAGTTAGTTCAACCAATCTGGGAAAT 1020
Db 961 TTTCTTGGGGGCTGATCAAGCAACATCTATAGTTAGTTCAACCAATCTGGGAAAT 1020
Qy 1021 AATGATGATGAAATCTCTCGGCTCCAAACCTTCCGCTCCAGAGAAATCTCCAGAG 1080
Db 1021 AATGATGATGAAATCTCTCGGCTCCAAACCTTCCGCTCCAGAGAAATCTCCAGAG 1080
Qy 1081 AACGTTGTCAGATATGTCACACCAATGCCATCTCTATGAGCGCTGGTGAACATCCC 1140
Db 1081 AACGTTGTCAGATATGTCACACCAATGCCATCTCTATGAGCGCTGGTGAACATCCC 1140
Qy 1141 GACCATGTTGTTGTTAATTAAGTATGTCCTTACGTAAGGGGATAGCAAGAGCCATGAT 1200
Db 1141 GACCATGTTGTTGTTAATTAAGTATGTCCTTACGTAAGGGGATAGCAAGAGCCATGAT 1200
Qy 1201 GAGTACACTTCAGAGATATTCATGAGGTGAGAAACCAATGTTTGGACCAACATATGT 1260
Db 1201 GAGTACACTTCAGAGATATTCATGAGGTGAGAAACCAATGTTTGGACCAACATATGT 1260
Qy 1261 GAGGATTCCTTTTACCTGCTCTATTAATCTTGAATGATGATCTTCTTGTGAGCTGAGC 1320
Db 1261 GAGGATTCCTTTTACCTGCTCTATTAATCTTGAATGATGATCTTCTTGTGAGCTGAGC 1320
Qy 1321 ACTAGAAATCCAGTTTAAAGCTGAAAAATGAGGAAAAATTCACATCATTTCCACCAATGCT 1380
Db 1321 ACTAGAAATCCAGTTTAAAGCTGAAAAATGAGGAAAAATTCACATCATTTCCACCAATGCT 1380
Qy 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTCTGTGTTCCACCGGGTACACCAAGTGTGAAT 1440
Db 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTCTGTGTTCCACCGGGTACACCAAGTGTGAAT 1440
```

```
Qy 1441 GCATTGTCAAAAGCAGCGTCGCAATGCTGAAAAATTAATGAGGCTTGTGTGATTTGCC 1500
Db 1441 GCATTGTCAAAAGCAGCGTCGCAATGCTGAAAAATTAATGAGGCTTGTGTGATTTGCC 1500
Qy 1501 CCAGGAATTAACATGATTTCTGAGTACAAAGTGA 1533
Db 1501 CCAGGAATTAACATGATTTCTGAGTACAAAGTGA 1533

RESULT 4
US-10-025-003-15
; Sequence 15, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-15

Query Match 99.9%; Score 1531.4; DB 14; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGTTCAATCGAATTTTAAGTAGAGAGTCTCTAATGTAAGTACACCGAGCTGAGATT 60
Db 1 ATGTTCAATCGAATTTTAAGTAGAGAGTCTCTAATGTAAGTACACCGAGCTGAGATT 60
Qy 61 CAGTCGCTGTACACATGAGAAACCAACGAACTTTGTTCAAGAGACAGAAATGSCATAT 120
Db 61 CAGTCGCTGTACACATGAGAAACCAACGAACTTTGTTCAAGAGACAGAAATGSCATAT 120
Qy 121 CAGTGAATTTGTAAACCCAAATCCGTCAACTACCAATTTAAACCAACCACTGTTCCA 180
Db 121 CAGTGAATTTGTAAACCCAAATCCGTCAACTACCAATTTAAACCAACCACTGTTCCA 180
Qy 181 AATTTGGGGGTATGCTTGTGGGTTGGGGTGAACCAACGCTTACCCTCAACCGGTGAT 240
Db 181 AATTTGGGGGTATGCTTGTGGGTTGGGGTGAACCAACGCTTACCCTCAACCGGTGAT 240
Qy 241 GTTATTTGTCACAGAGAGACATTTCAATGGGCTACAAAGACCAAGATTCAACAGCCAT 300
Db 241 GTTATTTGTCACAGAGAGACATTTCAATGGGCTACAAAGACCAAGATTCAACAGCCAT 300
Qy 301 TACTTGGCTCCCTCAACCAAGCTCAGCTATTCAGATTGATCTTCCAGGAGAGAGAA 360
Db 301 TACTTGGCTCCCTCAACCAAGCTCAGCTATTCAGATTGATCTTCCAGGAGAGAGAA 360
Qy 361 ATCTATGCCCATTCAGAGAGTCTGCTTCCATAGTGAATCTGACGACATGTGGTGGG 420
Db 361 ATCTATGCCCATTCAGAGAGTCTGCTTCCATAGTGAATCTGACGACATGTGTGGG 420
Qy 421 GGATGGGATATCAGCAACATGAACTTGGCTGATGCGCATGGCCAGGCAAAAGGTGTTGAC 480
Db 421 GGATGGGATATCAGCAACATGAACTTGGCTGATGCGCATGGCCAGGCAAAAGGTGTTGAC 480
Qy 481 ATCGATTTCCAGAGAGAGTTCAGGCGCTTACATGGAATTCATGGTTCCACTCCCGGGAATC 540
```



```
Db 481 ATGATTTGAGAGAGAGTTGAGGCTTATCATGGAATCCATGTTCCACTCCCGGAATC 540
Qy 541 TAGACCCCGGATTTCTGCTGCCAACAAGAGAGGCTCCAAACAAGTGAATGAAGGC 600
Db 541 TAGACCCCGGATTTCTGCTGCCAACAAGAGAGGCTCCAAACAAGTGAATGAAGGC 600
Qy 601 ACAAGCAAGAGAGAGTTGAGCAATCATCAAGAGAGTCAAGGCGTTTGAAGAGCACC 660
Db 601 ACAAGCAAGAGAGAGTTGAGCAATCATCAAGAGAGTCAAGGCGTTTGAAGAGCACC 660
Qy 661 AAAGTGACAGAGTGTGTCTCTGTGACCTGCCAACAAGAGAGTATAGCAATTTGTT 720
Db 661 AAAGTGACAGAGTGTGTCTCTGTGACCTGCCAACAAGAGAGTATAGCAATTTGTT 720
Qy 721 GTAGGCTTATATACACATGAGGAATCTTTGAGCTGTGTGACAGAGAAATGAGGCTGAG 780
Db 721 GTAGGCTTATATACACATGAGGAATCTTTGAGCTGTGTGACAGAGAAATGAGGCTGAG 780
Qy 781 ATTTCTCTTCCACCTTGTATGCGCATGCTGTGTGATGGAATAATGTTCTTTCATTAAT 840
Db 781 ATTTCTCTTCCACCTTGTATGCGCATGCTGTGTGATGGAATAATGTTCTTTCATTAAT 840
Qy 841 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATGATCTTGCATCGGAGAGAACACT 900
Db 841 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATGATCTTGCATCGGAGAGAACACT 900
Qy 901 TTGATTTGTGAGAGATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTTGAT 960
Db 901 TTGATTTGTGAGAGATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTTGAT 960
Qy 961 TTTCTGTGAGGCTGTGTATCAAGCAATCTATGTAATGTAACAACCTCTGGGAAC 1020
Db 961 TTTCTGTGAGGCTGTGTATCAAGCAATCTATGTAATGTAACAACCTCTGGGAAC 1020
Qy 1021 AATGATGATGATATCTCTGAGCTCCAAACCTTCCGCTCCAAAGAAATCTCCAAAGAG 1080
Db 1021 AATGATGATGATATCTCTGAGCTCCAAACCTTCCGCTCCAAAGAAATCTCCAAAGAG 1080
Qy 1081 AAGCTTTGTGAGAGATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTTGAT 1140
Db 1081 AAGCTTTGTGAGAGATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTTGAT 1140
Qy 1141 GACCATGTTGTGTTATTAAGTATGCGCTTACGTAGGAGATGAGCAAGAGCATGAT 1200
Db 1141 GACCATGTTGTGTTATTAAGTATGCGCTTACGTAGGAGATGAGCAAGAGCATGAT 1200
Qy 1201 GAGTACACTTCAAGAGATGATGAGTGAAGAAACCAATGTTTGTGACAAACATGT 1260
Db 1201 GAGTACACTTCAAGAGATGATGAGTGAAGAAACCAATGTTTGTGACAAACATGT 1260
Qy 1261 GAGGATTCCTTTTGTAGCTCTCTATTAATCTTGAATGATGCTTGTCTGTGAGTGAAG 1320
Db 1261 GAGGATTCCTTTTGTAGCTCTCTATTAATCTTGAATGATGCTTGTCTGTGAGTGAAG 1320
Qy 1321 ACTAGAAATCAGTTTAAAGCTGAAATGAGGAAATTTGACATCTTCCACCAAGTTGCT 1380
Db 1321 ACTAGAAATCAGTTTAAAGCTGAAATGAGGAAATTTGACATCTTCCACCAAGTTGCT 1380
Qy 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTGTGTTCCACCGGATACACAGTGTGAAT 1440
Db 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTGTGTTCCACCGGATACACAGTGTGAAT 1440
Qy 1441 GCATTTGTCAAGAGAGTGTGACCAAGGCTCTGTGTTCCACCGGATACACAGTGTGAAT 1500
Db 1441 GCATTTGTCAAGAGAGTGTGACCAAGGCTCTGTGTTCCACCGGATACACAGTGTGAAT 1500
Qy 1501 CGAGAGAAATACATGATTTCTGAGTACATGGA 1533
Db 1501 CGAGAGAAATACATGATTTCTGAGTACATGGA 1533
```

RESULT 5

```
US-10-718-952-13
; Sequence 13, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-13
```

```
Query Match 99.9%; Score 1531.4; DB 19; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ANGTTCATGAGAAATTTAAGTAGAGAGTCCATATGTAAGTACACCGAGACTGAGATT 60
Db 1 ANGTTCATGAGAAATTTAAGTAGAGAGTCCATATGTAAGTACACCGAGACTGAGATT 60
Qy 61 CAGTCCGTGTACACTACAGAAACCAACCGAATCTTGTACAGAGAACAGAAATGCACTAT 120
Db 61 CAGTCCGTGTACACTACAGAAACCAACCGAATCTTGTACAGAGAACAGAAATGCACTAT 120
Qy 121 CAGTGGATTTGTCAAAACCCAAATCCGTCACTACCAATTTTAAACCAACCCATGTTCCA 180
Db 121 CAGTGGATTTGTCAAAACCCAAATCCGTCACTACCAATTTTAAACCAACCCATGTTCCA 180
Qy 181 AATTTGGGGGTGATGCTTGTGGGGTGGGGTGAACCAAGGCTTACCCCTCAACCGGTGAT 240
Db 181 AATTTGGGGGTGATGCTTGTGGGGTGGGGTGAACCAAGGCTTACCCCTCAACCGGTGAT 240
Qy 241 GTTATTTCTAACAGAGAGACATTTCAATGAGCTTACAAAGAGACAAAGATTTCAACAGCCAAAT 300
Db 241 GTTATTTCTAACAGAGAGACATTTCAATGAGCTTACAAAGAGACAAAGATTTCAACAGCCAAAT 300
Qy 301 TACTTTGGCTCCCTCAACCCCAAGCTTCACTATTTCCAGATTGATCTTCCAGGAGAGAA 360
Db 301 TACTTTGGCTCCCTCAACCCCAAGCTTCACTATTTCCAGATTGATCTTCCAGGAGAGAA 360
Qy 361 ACTATAGCCCAATTAAGAGTCTGCTTCCAAAGTGAATTCCTGACGACATGTTGTTGGG 420
Db 361 ACTATAGCCCAATTAAGAGTCTGCTTCCAAAGTGAATTCCTGACGACATGTTGTTGGG 420
Qy 421 GGATGGATATCAGAACATGAACTGCTGTATGTCATGCGCAAGGCAAGGAGTGTGAGC 480
Db 421 GGATGGATATCAGAACATGAACTGCTGTATGTCATGCGCAAGGCAAGGAGTGTGAGC 480
Qy 481 ATGCAATTTGACAGAGAGTGTGAGGCTTATCAGAAATCATGAGTTCACCTCCCGGAATC 540
Db 481 ATGCAATTTGACAGAGAGTGTGAGGCTTATCAGAAATCATGAGTTCACCTCCCGGAATC 540
Qy 541 TAGGACCCGGATTTCAATGCTGCAACAGAGAGCGTGCCAAACAGTGAATTAAGGAGC 600
Db 541 TAGGACCCGGATTTCAATGCTGCAACAGAGAGCGTGCCAAACAGTGAATTAAGGAGC 600
Qy 601 ACAAGCAAGAGAGATTTCAAGAAATCATCAAGAGATCAATCAAGGCTTTAAGAGAGCACC 660
Db 601 ACAAGCAAGAGAGATTTCAAGAAATCATCAAGAGATCAATCAAGGCTTTAAGAGAGCACC 660
```

QY 661 AAAGTGACAAAGTGGTGTCTCTGTGACTGCGCAACACAGAGAGGTATAGCAATTTGGT 720
DB 661 AAAGTGACAAAGTGGTGTCTCTGTGACTGCGCAACACAGAGAGGTATAGCAATTTGGT 720
QY 721 GTAGGCTTTAATGACACATGAGAAATCTCTGGCTGCTGTGACAGAAATAGAGCTGAG 780
DB 721 GTAGGCTTTAATGACACATGAGAAATCTCTGGCTGCTGTGACAGAAATAGAGCTGAG 780
QY 781 ATTTCTCTTCCACCTGTATGCAATGCTGTGATGAGAAATGTTCTTTCATTAAT 840
DB 781 ATTTCTCTTCCACCTGTATGCAATGCTGTGATGAGAAATGTTCTTTCATTAAT 840
QY 841 GGAGAGCCCTCAGAGACATTTTGTACAGAGGCTGATTTGATCTTCCATGCGAGAACT 900
DB 841 GGAGAGCCCTCAGAGACATTTTGTACAGAGGCTGATTTGATCTTCCATGCGAGAACT 900
QY 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGTTAT 960
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGTTAT 960
QY 961 TTTCTTGTGGGGGCTGTATCAAGCCCATGATATGTTAACAACCATGCGGGAAC 1020
DB 961 TTTCTTGTGGGGGCTGTATCAAGCCCATGATATGTTAACAACCATGCGGGAAC 1020
QY 1021 AATGATGTATGATCTCTCGGCTCCACAACCTTCCGCTCAGAGAAATCTTCAAGAGC 1080
DB 1021 AATGATGTATGATCTCTCGGCTCCACAACCTTCCGCTCAGAGAAATCTTCAAGAGC 1080
QY 1081 AAGCTTTGTGACGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGTAACATCCC 1140
DB 1081 AAGCTTTGTGACGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGTAACATCCC 1140
QY 1141 GACCATGTTGTTATTAAGATATGACCTTACGTAGGGGATAGCAAGAGAGCATGAT 1200
DB 1141 GACCATGTTGTTATTAAGATATGACCTTACGTAGGGGATAGCAAGAGAGCATGAT 1200
QY 1201 GAGTACACTTCAAGATATTCATGAGTGAAGAACACCATTTGTTGACACACATGT 1260
DB 1201 GAGTACACTTCAAGATATTCATGAGTGAAGAACACCATTTGTTGACACACATGT 1260
QY 1261 GAGGATTCCTTTTACCTGCTCTATTAATCTTGAATCTTGGTCTTGTGTGAGCTGAGC 1320
DB 1261 GAGGATTCCTTTTACCTGCTCTATTAATCTTGAATCTTGGTCTTGTGTGAGCTGAGC 1320
QY 1321 ACTAGATTCAGTTTAAAGCTGAAGTGAAGGAGAAATTCACATCTTCCAGCCAGTTGCT 1380
DB 1321 ACTAGATTCAGTTTAAAGCTGAAGTGAAGGAGAAATTCACATCTTCCAGCCAGTTGCT 1380
QY 1381 ACCATTCAGCTATCTGACCAAGGCTCTCTGTGTTCAACCGGGTACACCAAGTGTGAAT 1440
DB 1381 ACCATTCAGCTATCTGACCAAGGCTCTCTGTGTTCAACCGGGTACACCAAGTGTGAAT 1440
QY 1441 GCATTGTCAAGAGAGCTGCAATGCTGGAAGAAATATGAGGCTTGTGTGATTTGGCC 1500
DB 1441 GCATTGTCAAGAGAGCTGCAATGCTGGAAGAAATATGAGGCTTGTGTGATTTGGCC 1500
QY 1501 CCAGAGATPAACATGATCTCGAGTACAGTGA 1533
DB 1501 CCAGAGATPAACATGATCTCGAGTACAGTGA 1533

RESULT 6

US-10-718-952-15
; Sequence 15, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitez, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C

; CURRENT APPLICATION NUMBER: US/10/718, 952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-718-952-15

Query Match 99.9%; Score 1531.4; DB 19; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTCAATCAGAAATTTAAGGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
DB 1 ATGTTCAATCAGAAATTTAAGGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
QY 61 CAGTCCGTGTACAACTACGAAACCAACCGAATCTTGTTCAGAGAGAGAGGACCTAT 120
DB 61 CAGTCCGTGTACAACTACGAAACCAACCGAATCTTGTTCAGAGAGAGAGGACCTAT 120
QY 121 CAGTGAATGTCAAAACCCAAATCCGTCAACTCAATTTAAACCAACCAACCATGTTCA 180
DB 121 CAGTGAATGTCAAAACCCAAATCCGTCAACTCAATTTAAACCAACCAACCATGTTCA 180
QY 181 AAATGGGGGTGATCTGTGGGTTGGGGTGAAGAACAAACGCTCTACCTCACCCTGGT 240
DB 181 AAATGGGGGTGATCTGTGGGTTGGGGTGAAGAACAAACGCTCTACCTCACCCTGGT 240
QY 241 GTTATTCCTAAGAGAGAGACATTCATGAGGCTACAAAGAGAGCAAGATTCAACAGCAAT 300
DB 241 GTTATTCCTAAGAGAGAGAGACATTCATGAGGCTACAAAGAGAGCAAGATTCAACAGCAAT 300
QY 301 TACTTGGCTCCCTCAGCCCAACGCTCAGCTATTCGAGTTGATTCCTTCCAGGAGAGAA 360
DB 301 TACTTGGCTCCCTCAGCCCAACGCTCAGCTATTCGAGTTGATTCCTTCCAGGAGAGAA 360
QY 361 ATCTATGCCCATTCAGAGTCTGCTTCCATAGTTTAATCTTGAAGCAATTTGTGTTGGG 420
DB 361 ATCTATGCCCATTCAGAGTCTGCTTCCATAGTTTAATCTTGAAGCAATTTGTGTTGGG 420
QY 421 GATGGGATATGAGCAAGTGAACCTGCTGATGTCATGCGCAGGAGCAAGGTTGTGAC 480
DB 421 GATGGGATATGAGCAAGTGAACCTGCTGATGTCATGCGCAGGAGCAAGGTTGTGAC 480
QY 481 ATCGATTTGCAAGAGCAGTTGAGGCTTACATGGAATCCATGATTCCTCCCGGAATC 540
DB 481 ATCGATTTGCAAGAGCAGTTGAGGCTTACATGGAATCCATGATTCCTCCCGGAATC 540
QY 541 TACGACCCGGAATTCATGCTGTGCAACAGAGAGAGCTGCCAACAGCTGATTAAGGGC 600
DB 541 TACGACCCGGAATTCATGCTGTGCAACAGAGAGAGCTGCCAACAGCTGATTAAGGGC 600
QY 601 ACAAGCAAGACCAAGTTGAGCAAAATCATCAAGACATCAAGGCTTTAAGAGGACACC 660
DB 601 ACAAGCAAGACCAAGTTGAGCAAAATCATCAAGACATCAAGGCTTTAAGAGGACACC 660
QY 661 AAAGTGACAAAGTGGTGTCTCTGTGACTGCGCAACACAGAGAGGTATAGCAATTTGGTT 720
DB 661 AAAGTGACAAAGTGGTGTCTCTGTGACTGCGCAACACAGAGAGGTATAGCAATTTGGTT 720
QY 721 GTAGGCTTTAATGACACATGAGAAATCTCTGGCTGCTGTGACAGAAATAGAGCTGAG 780
DB 721 GTAGGCTTTAATGACACATGAGAAATCTCTGGCTGCTGTGACAGAAATAGAGCTGAG 780
QY 781 ATTTCTCTTCCACCTGTGATGCAATGCTGTGATGAGAAATGTTCTTTCATTAAT 840
DB 781 ATTTCTCTTCCACCTGTGATGCAATGCTGTGATGAGAAATGTTCTTTCATTAAT 840

```

QY      841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTGATCTTGCATCGGAGGAACACT 900
      |||
Db      841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTGATCTTGCATCGGAGGAACACT 900
QY      901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTTGTTGAT 960
      |||
Db      901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTTGTTGAT 960
QY      961 TTTCTTTGAGGAGCTGTGATCAAGCCACATCTATAGTTATCAACCATCTGGGAAAC 1020
      |||
Db      961 TTTCTTTGAGGAGCTGTGATCAAGCCACATCTATAGTTATCAACCATCTGGGAAAC 1020
QY      1021 AATGATGATGATGATCTCTCGGCTCCAAACCTTCCCTCCAGAGAAATCTTCAAGAGC 1080
      |||
Db      1021 AATGATGATGATGATCTCTCGGCTCCAAACCTTCCCTCCAGAGAAATCTTCAAGAGC 1080
QY      1081 AACGTTGTTAGCAATATGATCAACAGCAATGCCATCCTCTATGAGGCTGGTGAACATCCC 1140
      |||
Db      1081 AACGTTGTTAGCAATATGATCAACAGCAATGCCATCCTCTATGAGGCTGGTGAACATCCC 1140
QY      1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGAGGATGACAAAGAGCCATGAT 1200
      |||
Db      1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGAGGATGACAAAGAGCCATGAT 1200
QY      1201 GAGTACACTTCAGAGATATTCATGAGTGAAGAAACACATGTTTTCACACACATGT 1260
      |||
Db      1201 GAGTACACTTCAGAGATATTCATGAGTGAAGAAACACATGTTTTCACACACATGT 1260
QY      1261 GAGATTCCTCTTTTAACTGCTCCTATTATCTTGAATGCTTCTTCTGCTGAGTAC 1320
      |||
Db      1261 GAGATTCCTCTTTTAACTGCTCCTATTATCTTGAATGCTTCTTCTGCTGAGTAC 1320
QY      1321 ACTAGATACAGTTTAACTGAACTGAAATGAGGAAATTCACATTCACACCATGTCCT 1380
      |||
Db      1321 ACTAGATACAGTTTAACTGAACTGAAATGAGGAAATTCACATTCACACCATGTCCT 1380
QY      1381 ACCATTCTCAGCTATATGACCAAGGCTCCTCTGCTTCCACCGGATCACCAAGTGTGAT 1440
      |||
Db      1381 ACCATTCTCAGCTATATGACCAAGGCTCCTCTGCTTCCACCGGATCACCAAGTGTGAT 1440
QY      1441 GCATTGTCAAGCAGCGTGCATATGCTGAAACATATGAGGCTGTGTGTTGATTGGCC 1500
      |||
Db      1441 GCATTGTCAAGCAGCGTGCATATGCTGAAACATATGAGGCTGTGTGTTGATTGGCC 1500
QY      1501 CCAGAGATATACATGATTCCTCGAGTACAGTGA 1533
      |||
Db      1501 CCAGAGATATACATGATTCCTCGAGTACAGTGA 1533

```

```

; ORGANISM: Glycine max
US-10-025-003-9
Query Match      95.5%; Score 1464.2; DB 14; Length 1533;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY      1 ATGTTTATGAGAAATTTTAAAGTAAAGTCTCTAATATGAAATATACCCGAGCTGAGATT 60
      |||
Db      1 ATGTTTATGAGAAATTTTAAAGTAAAGTCTCTAATATGAAATATACCCGAGCTGAGATT 60
QY      61 CAGTCCGTGTACAACTACGAAACCAACCGAATCTTGTTCAGAGAAACAGAAATGGCACTTAT 120
      |||
Db      61 CAGTCCGTGTACAACTACGAAACCAACCGAATCTTGTTCAGAGAAACAGAAATGGCACTTAT 120
QY      121 CAGTGGATTGTCAAAACCAAAATCCGTCAACTCAATTTTAAACCAACCAATGTTTCA 180
      |||
Db      121 CAGTGGATTGTCAAAACCAAAATCCGTCAACTCAATTTTAAACCAACCAATGTTTCA 180
QY      181 AAATTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCTTACCGGTGGT 240
      |||
Db      181 AAATTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCTTACCGGTGGT 240
QY      241 GTTATTTGCTAACAGAGAGACATTTCAATGAGCTAACAAAGACAAAGATTCACAGCCAAAT 300
      |||
Db      241 GTTATTTGCTAACAGAGAGACATTTCAATGAGCTAACAAAGACAAAGATTCACAGCCAAAT 300
QY      301 TACTTTGGCTCCCTACCCCAAGCTTCAAGTATTCAGATTTGATCTTCCAGGAGAGAA 360
      |||
Db      301 TACTTTGGCTCCCTACCCCAAGCTTCAAGTATTCAGATTTGATCTTCCAGGAGAGAA 360
QY      361 ATCTATGCCCCATTCAAGAGTCTGCTTCCAAATGTTAAATCTTGAAGACATTTGTGTTGGG 420
      |||
Db      361 ATCTATGCCCCATTCAAGAGTCTGCTTCCAAATGTTAAATCTTGAAGACATTTGTGTTGGG 420
QY      421 GGATGGGATATACAGCAACATGAACTGAGTATGCAATGAGGACAGGACAAAGGTTTGAAC 480
      |||
Db      421 GGATGGGATATACAGCAACATGAACTGAGTATGCAATGAGGACAGGACAAAGGTTTGAAC 480
QY      481 ATCGATTTGACAGAGCAGTTGAGGCTTATCATGAAATCCATGTTTCACTCCCGGAATC 540
      |||
Db      481 ATCGATTTGACAGAGCAGTTGAGGCTTATCATGAAATCCATGTTTCACTCCCGGAATC 540
QY      541 TACGACCCGAGATTTCATGCTGCTCCCAACAGAGAGCGTGCACACAGCTGATTAAGGCG 600
      |||
Db      541 TACGACCCGAGATTTCATGCTGCTCCCAACAGAGAGCGTGCACACAGCTGATTAAGGCG 600
QY      541 TATGACCCGAGATTTCATGCTGCTCCCAACAGAGAGCGTGCACACAGCTGATTAAGGCG 600
      |||
Db      541 TATGACCCGAGATTTCATGCTGCTCCCAACAGAGAGCGTGCACACAGCTGATTAAGGCG 600
QY      601 ACAAGACAAAGACAGTTTCAGCAATTCATCAAGACATCAAGAGCGTTTAAAGAGCACCC 660
      |||
Db      601 ACAAGACAAAGACAGTTTCAGCAATTCATCAAGACATCAAGAGCGTTTAAAGAGCACCC 660
QY      661 AAAGTGAACAAGGTGTTGCTCTGTGACTGCCAACACAGAGGTATAGCAATTTGTT 720
      |||
Db      661 AAAGTGAACAAGGTGTTGCTCTGTGACTGCCAACACAGAGGTATAGCAATTTGTT 720
QY      721 GTAGGCTTTAATGACCAATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
      |||
Db      721 GTAGGCTTTAATGACCAATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
QY      781 ATTTTCCCTTCCACCTGTGATGCAATGCTGCTGTGATGAGAAATGTTTCTTCAATTAAT 840
      |||
Db      781 ATTTTCCCTTCCACCTGTGATGCAATGCTGCTGTGATGAGAAATGTTTCTTCAATTAAT 840
QY      841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTGATCTTGCATCGGAGGAACACT 900
      |||
Db      841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTGATCTTGCATCGGAGGAACACT 900
QY      901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTTGTTGAT 960
      |||
Db      901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTTGTTGAT 960
QY      961 TTTCTTTGAGGAGCTGTGATCAAGCCACATCTATAGTTATCAACCATCTGGGAAAC 1020
      |||

```

Db 961 TTCCTTGCGGGGCTGTATCAAGCCAACTATAGTCAGTTACCAACCATCTGGGAAAC 1020
Qy 1021 AATGATGATGAATCTCTGGGCTCCAGAACCTCCGCTCCAGGAATCTCCAGAGC 1080
Db 1021 AATGATGATGAATCTCTGGGCTCCAGAACCTCCGCTCCAGGAATCTCCAGAGC 1080
Qy 1081 AACGTTGTGACGATATGCTCAACAGCAATGCAATCTCTATAGAGCTGTGAAATCCC 1140
Db 1081 AACGTTGTGATATGCTCAACAGCAATGCAATCTCTATAGAGCTGTGAAATCCA 1140
Qy 1141 GACCATGTTGTTATTAATGATGCTTACGTAAGGAGATGCAAGAGACCATGAT 1200
Db 1141 GACCATGTTGTTATTAATGATGCTTACGTAAGGAGATGCAAGAGACCATGAT 1200
Qy 1201 GAGTCACTTCAAGATATTCATGGGTGGAAGAACCATTTGTTTGCAACACATGT 1260
Db 1201 GAGTCACTTCAAGATATTCATGGGTGGAAGAACCATTTGTTTGCAACACATGC 1260
Qy 1261 GAGCATCCCTTTTAACTGCTCTATTAATCTTGACCTTCTTCTGCTGAGCTGAC 1320
Db 1261 GAGCATCCCTCTTACTGCTCTCTATTAATCTTGACCTTCTTCTGCTGAGCTGAC 1320
Qy 1321 ACTAGATCAGTTTAACTGAAATGAGGAAATTCACACTCATTCACCCAGTTGCT 1380
Db 1321 ACTAGATCAGTTTAACTGAAATGAGGAAATTCACACTCATTCACCCAGTTGCT 1380
Qy 1381 ACCATTCAGCATCTGACCAAGGCTCTGTTTCCACCGGATACACAGTGTGAT 1440
Db 1381 ACCATTCAGCATCTGACCAAGGCTCTGTTTCCACCGGATACACAGTGTGAT 1440
Qy 1441 GCATTGTCAAGCAGGTCGAATGCTGAAATCATATGAGGCTTGTGTGATTTGGCC 1500
Db 1441 GCATTGTCAAGCAGGTCGAATGCTGAAATCATATGAGGCTTGTGTGATTTGGCC 1500
Qy 1501 CCAGAGATTAACATGATTTCTCGAGTACCAAGTGA 1533
Db 1501 CCAGAGATTAACATGATTTCTCGAGTACCAAGTGA 1533

RESULT 8

US-10-718-952-9
; Sequence 9, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAPEINASE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-9

Query Match 95.5%; Score 1464.2; DB 19; Length 1533;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 ATGTTATGAGAAATTTTAAAGTAGAGAGTCCCTAATAGTGAAGTACCCGAGACTGAGATT 60
Db 1 ATGTTATGAGAAATTTTAAAGTAGAGTGTCTTAATAGTGAAGTACCCGAGACTGAGATT 60

Qy 61 CAGTCCGTATCAACTAGAAAACCCAGACTTGTTCAGAGAAAGAGAAATGGCACTAT 120
Db 61 CAGTCCGTATCAACTAGAAAACCCAGACTTGTTCAGAGAAAGAGAAATGGCACTAT 120
Qy 121 CAGTGGATTTGCAAAACCCAAATCCGTCAATACAAATTTAAACCAACCAATGTTCCA 180
Db 121 CAGTGGATTTGCAAAACCCAAATCCGTCAATACAAATTTAAACCAACCAATGTTCCCT 180
Qy 181 AAATGGGGGATGATCTTGTGGGTTGGGGTGAACCAACGCGCTACCGCTACCGGTGT 240
Db 181 AAATGGGGGATGATCTTGTGGGTTGGGGTGAACCAACGCGCTACCGGTGT 240
Qy 241 GTTATGCTTAAAGAGAGAACATTTCAATGGGCTACAAAGAGACAGATTTCAACAGCAAT 300
Db 241 GTTATGCTTAAAGAGAGAACATTTCAATGGGCTACAAAGAGACAGATTTCAACAGCAAT 300
Qy 301 TACTTTGGCTCCCTACCCAGAGCTCAGCTATTCGAGTTGATTCCTTCAAGGAGAGGA 360
Db 301 TACTTTGGCTCCCTACCCAGAGCTCAGCTATTCGAGTTGATTCCTTCAAGGAGAGGA 360
Qy 361 ATCTATGCCCATTTCAAGAGTCTGCTTCAATGTTAATCCTGACGACATTTGTTGGG 420
Db 361 ATCTATGCCCATTTCAAGAGTCTGCTTCAATGTTAATCCTGACGACATTTGTTGGG 420
Qy 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCTATGGCCAGGCAAAAGTGTGAC 480
Db 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCTATGGCCAGGCAAAAGTGTGAC 480
Qy 481 ATCGATTTGCAAGACAGTGAAGGCTTTCATGAAATCAGTGTTCCTCCCGGGAATC 540
Db 481 ATCGATTTGCAAGACAGTGAAGGCTTTCATGAAATCAGTGTTCCTCCCGGGAATC 540
Qy 541 TACGACCCGGAATTTATTGCTGCTCCAAACCAAGAGAGCGTGCCAAACAGTGTAAAGGC 600
Db 541 TATGACCCGGAATTTATTGCTGCTCCAAACCAAGAGAGCGTGCCAAACAGTGTAAAGGC 600
Qy 601 ACAAAGCAAGAGCAAGTTCAGCAATCATCAAGATCATCAAGCGCTTTAAGAACCCACC 660
Db 601 ACAAAGCAAGAGCAAGTTCAGCAATCATCAAGATCATCAAGCGCTTTAAGAACCCACC 660
Qy 661 AAAGTGACAAAGTGTGCTGCTGACCTGCAACAGAGAGTATTAAGCAATTTGGTT 720
Db 661 AAAGTGACAAAGTGTGCTGCTGACCTGCAACAGAGAGTATTAAGCAATTTGGTT 720
Qy 721 GTAGGCTTAAATGACCAATGAGAGATCTCTGGGCTGCTGTGAGACAGAAATGAGGCTGAG 780
Db 721 GTAGGCTTAAATGACCAATGAGAGATCTCTGGGCTGCTGTGAGACAGAAATGAGGCTGAG 780
Qy 781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGATGAAATGTTCTTCAATTAAT 840
Db 781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGATGAAATGTTCTTCAATTAAT 840
Qy 841 GGAAGCCCTCAGAAACATTTTGTACAGAGGCTGATTTGATCTTGCATGCGGAGAACCT 900
Db 841 GGAAGCCCTCAGAAACATTTTGTACAGAGGCTGATTTGATCTTGCATGCGGAGAACCT 900
Qy 901 TTGATTTGGTGAATGACTTCAAGAGTGTGACCAACCAATGAAATCTGTGTGTTGAT 960
Db 901 TTGATTTGGTGAATGACTTCAAGAGTGTGACCAACCAATGAAATCTGTGTGTTGAT 960
Qy 961 TTTCTTGTGGGGGCTGTATCAAGCCAACTATATAGTTAGTTACCAACCATCTGGGAAAC 1020
Db 961 TTTCTTGTGGGGGCTGTATCAAGCCAACTATATAGTTAGTTACCAACCATCTGGGAAAC 1020
Qy 1021 AATGATGATGAATCTCTGGGCTCCAGAACCTTCCGCTCCAGGAATCTCCAGAGC 1080
Db 1021 AATGATGATGAATCTCTGGGCTCCAGAACCTTCCGCTCCAGGAATCTCCAGAGC 1080
Qy 1081 AACGTTGTGACGATATGCTCAACAGCAATGCAATCTCTATAGAGCTGTGAAATCCC 1140
Db 1081 AACGTTGTGATATGCTCAACAGCAATGCAATCTCTATAGAGCTGTGAAATCCA 1140
Qy 1141 GACCATGTTGTTATTAATGATGCTTACGTAAGGAGATGACAGAGACCATGAT 1200

Db 1141 GACCACTGTTGTTTATTAAGTATGTCCTTACGTAGGGACAGCAAGAGCCATGAT 1200
Qy 1201 GAGTACACTTCAGAGATATTATCATGGGTGGAAGAACCACTGTTTTGCAACAACATGTT 1260
Db 1201 GAGTACACTTCAGAGATATTATCATGGGTGGAAGAACCACTGTTTTGCAACAACATGTT 1260
Qy 1261 GAGGATTCCTTTTACTGCTCTATTTATTTTGGACTTGGTCTTTCTTGTGAGCTGAGC 1320
Db 1261 GAGGATTCCTTTTACTGCTCTATTTATTTTGGACTTGGTCTTTCTTGTGAGCTGAGC 1320
Qy 1321 ACTAGATTCAGATTAAAGCTGAAGTGAAGGAAAAATTCACATCTATTCACCCAGTTGCT 1380
Db 1321 ACTAGATTCAGATTAAAGCTGAAGTGAAGGAAAAATTCACATCTATTCACCCAGTTGCT 1380
Qy 1381 ACCATTCTAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACACCAAGTGTGAAT 1440
Db 1381 ACCATTCTAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACACCAAGTGTGAAT 1440
Qy 1441 GCATTGTCAAAGAGCGGTGCAATGCTGGAAGAAATATGAGGGCTTGTGTGATTGGCC 1500
Db 1441 GCATTGTCAAAGAGCGGTGCAATGCTGGAAGAAATATGAGGGCTTGTGTGATTGGCC 1500
Qy 1501 CCAGAGATTAACATGATTCCTGAGTACCAAGTGA 1533
Db 1501 CCAGAGATTAACATGATTCCTGAGTACCAAGTGA 1533

RESULT 9

US-10-025-003-1
; Sequence 1, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-1

Query Match 95.5%; Score 1464.2; DB 14; Length 1760;

Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 ATGTTATCGAGAAATTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 60
Db 54 ATGTTATCGAGAAATTTAAAGTAGAGTCTTAATGTGAAGTACACCGAGCTGAGATT 113
Qy 61 CAGTCCGTGTACAACTAGAAACACCGAATTGTTACGAGAAACAGAAATGGACCTAT 120
Db 114 CAGTCCGTGTACAACTAGAAACACCGAATTGTTACGAGAAACAGAAATGGACCTAT 173
Qy 121 CAGTGAATTTGTAACCAACCAATCCGTCACTACATTTTAAACCAACCACTGTTTCA 180
Db 174 CAGTGAATTTGTAACCAACCAATCCGTCACTACATTTTAAACCAACCACTGTTTCT 233
Qy 181 AATTTGGGGGTGATCTTGTGGGTGGAACCAAGCTCTACCCCTCAACGGGTGAT 240
Db 234 AATTTGGGGGTGATCTTGTGGGTGGAACCAAGCTCTACCCCTCAACGGGTGAT 293

Qy 241 GTTATTGCTTAACAGAGAGCAATTTCAATGGGCTACAAAGAGCAAGATTCAACAGCCAT 300
Db 294 GTTATTGCTTAACAGAGAGCAATTTCAATGGGCTACAAAGAGCAAGATTCAACAGCCAT 353
Qy 301 TACTTTGGCTCCCTCAACCAACCTCAGCTATTTCAGTTGATTCCTTCCAGGAGAGAA 360
Db 354 TACTTTGGCTCCCTCAACCAACCTCAGCTATTTCAGTTGATTCCTTCCAGGAGAGAA 413
Qy 361 ATCTATGCCCATTTCAAGAGTGTGCTTCCATGATTTAATCTGACGACATTTGTTGGG 420
Db 414 ATCTATGCCCATTTCAAGAGTGTGCTTCCATGATTTAATCTGACGACATTTGTTGGG 473
Qy 421 GGATGGATTAATACGACAACTGAACCTGCTATGCTATGCTGCAAGGCAAAAGTGTGAC 480
Db 474 GGATGGATTAATACGACAACTGAACCTGCTATGCTATGCTGCAAGGCAAAAGTGTGAC 533
Qy 481 ATCGATTTGACAGAGAGATGAGGCTTACATGGAATCCATGATTCACCTCCCGGAATC 540
Db 534 ATCGATTTGACAGAGAGATGAGGCTTACATGGAATCCATGATTCACCTCCCGGAATC 593
Qy 541 TACGACCCGGAATTTCAATGCTGCTCAACCAAGAGAGCTGCAACACCTGATTAGAGGC 600
Db 594 TATGACCCGGAATTTCAATGCTGCTCAACCAAGAGAGCTGCAACACCTGATTAGAGGC 653
Qy 601 ACMAAGCAAGCAAGTTCAGCAATCATCAAAAGCATTAAGGCTTTTAAAGAGCAGCC 660
Db 654 ACMAAGCAAGCAAGTTCMAAATCATCAAAAGCATCAAGGCTTTTAAAGAGCAGCC 713
Qy 661 AAAGTGAACAAGTGTGCTGCTGAGACTGCAACACAGAGAGGATTAAGCAATTTGGTT 720
Db 714 AAAGTGAACAAGTGTGCTGAGACTGCAACACAGAGAGGATTAAGCAATTTGGTT 773
Qy 721 GTAGGCTTTAATGACACATGAGAGATCTTTGCTGCTGTGAGACAGAAATGAGCTGAG 780
Db 774 GTAGGCTTTAATGACACATGAGAGATCTTTGCTGCTGTGAGACAGAAATGAGCTGAG 833
Qy 781 ATTTCTCTTCCACTGTTATGCTGCTGCTGTGATGAGAAATGTTCTTTCAATAT 840
Db 834 ATTTCTCTTCCACTGTTATGCTGCTGCTGTGATGAGAAATGTTCTTTCAATAT 893
Qy 841 GGAAGCCCTCAGAACCTTTGTACAGGGGCTGATGATCTTGCATGCGAGGAACAT 900
Db 894 GGAAGCCCTCAGAACCTTTGTACAGGGGCTGATGATCTTGCATGCGAGGAACAT 953
Qy 901 TTGATTTGAGAGATGACTTCAAGAGTGTGACCAACCAATGAATCTGTGTTGAT 960
Db 954 TTGATTTGAGAGATGACTTCAAGAGTGTGACCAACCAATGAATCTGTGTTGAT 1013
Qy 961 TTTCTTTGGGGGCTGTATCAAGCCAACATCTATATGTTAGTTACCAACCATTTGGGAAC 1020
Db 1014 TTTCTTTGGGGGCTGTATCAAGCCAACATCTATATGTTAGTTACCAACCATTTGGGAAC 1073
Qy 1021 AATGATGTGATGAATCTCCGGCTCCACAAACCTCCGTCGAAGAAATCTCCAAAGGC 1080
Db 1074 AATGATGTGATGAATCTCCGGCTCCACAAACCTCCGTCGAAGAAATCTCCAAAGGC 1133
Qy 1081 AACGTTGTTGAGATGATGCTCAACAGCAATGCTCTATGAGCTGTGATGATCC 1140
Db 1134 AACGTTGTTGAGATGATGCTCAACAGCAATGCTCTATGAGCTGTGATGATCC 1193
Qy 1141 GACCAATGTTGTTATTAATATGATGTGCTTAAGTGAAGGAGATGACAGACCATGAT 1200
Db 1194 GACCAATGTTGTTATTAATATGATGTGCTTAAGTGAAGGAGATGACAGACCATGAT 1253
Qy 1201 GAGTACACTTCAGATATATTCATGGGTGGAAGAACACATGTTTGTGACAAACATGT 1260
Db 1254 GAGTACACTTCAGATATATTCATGGGTGGAAGAACACATGTTTGTGACAAACATGT 1313
Qy 1261 GAGATTCCTCTTTTGTGCTCTTATATCTTGAATTTGATCTTGTGAGCTGAGC 1320
Db 1314 GAGATTCCTCTTTTGTGCTCTTATATCTTGAATTTGATCTTGTGAGCTGAGC 1373

```
QY 1321 ACTAGATTCAGTTTAAAGCTGAAAATGAGGGAAAAATTCACATTCACCCAGTTGCT 1380
| | | | |
DB 1374 ACTAGATTCAGTTTAAAGCTGAAAATGAGGGAAAAATTCACATTCACCCAGTTGCT 1433
| | | | |
QY 1381 ACCATTCTCAGCTATCTGACCAAGGCTCCTCTGTTCCACCGGTCACACATGTGGAT 1440
| | | | |
DB 1434 ACCATCTCAGCTATCTGACCAAGGCTCCTCTGTTCCACCGGTCACACATGTGGAT 1493
| | | | |
QY 1441 GCATTCTCAAGAGCGTCAATGCTGAAAACATATAGAGGCTGTGTGGATTGGCC 1500
| | | | |
DB 1494 GCATTCTCAAGAGCGTCAATGCTGAAAACATATAGAGGCTGTGTGGATTGGCC 1553
| | | | |
QY 1501 CCAGAGATPAACATGATTTCTCGATCAAGTGA 1533
| | | | |
DB 1554 CCAGAGATPAACATGATTTCTCGATCAAGTGA 1586
| | | | |

RESULT 10
US-10-718-952-1
; Sequence 1, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BR-1077-C
; CURRENT APPLICATION NUMBER: US/10/718, 952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-1

Query Match 95.5%; Score 1464.2; DB 19; Length 1760;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 ANCTTATGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAATGACCGAGCTGAGAT 60
| | | | |
DB 54 ATGTTATCGAGAAATTTTAAAGTAGAGTCTTAATGTGAATGACCGAGCTGAGAT 113
| | | | |
QY 61 CAGTCCGTGTAACAACGAAACACCGAATCTGTACGAGAACAGAAATGGACCTAT 120
| | | | |
DB 114 CAGTCCGTGTAACAACGAAACACCGAATCTGTGTTCACGAGAACAGAAATGGACCTAT 173
| | | | |
QY 121 CAGTGAATGTCAAAACCAATCCGTCAATCAATTTTAAACCAACACCAATGTTCA 180
| | | | |
DB 174 CAGTGAATGTCAAAACCAATCCGTCAATCAATTTTAAACCAACCAATGTTTCT 233
| | | | |
QY 181 AAATTTGGGGGTATGCTTGTGGGTGGGGTGAACAAAGGCTTACCTCAACCGTGT 240
| | | | |
DB 234 AAATTTGGGGGTATGCTTGTGGGTGGGGTGAACAAAGGCTTACCTCAACCGTGT 293
| | | | |
QY 241 GTTATGCTAAGAGAGGACATTTATGAGGCTTACAAAGACAAAGATTCAAAGCAAT 300
| | | | |
DB 294 GTTATGCTAAGAGAGGACATTTATGAGGCTTACAAAGACAAAGATTCAAAGCAAT 353
| | | | |
QY 301 TACTTTGGCTCCCTCAAGGCTCAGCTATTCAGATTGATCTTCCAGGGAGAGAA 360
| | | | |
DB 354 TACTTTGGCTCCCTCAAGGCTCAGCTATTCAGATTGATCTTCCAGGGAGAGAA 413
| | | | |
QY 361 ATCTATGCCCATTCAGAGTCTGCTTCCAATGTTAATCTGACACATTTGTTGGG 420
| | | | |
```

```
DB 414 ATCTATGCCCATTCAGAGCTGCTTCCAATGTTAATCCTGACGACATTTGTTGGG 473
| | | | |
QY 421 GGATGGATATCAGCAATGAACCTGCTGATGTCATGCGCAGGGCAAGATGTTTAC 480
| | | | |
DB 474 GGATGGATATCAGCAATGAACCTGCTGATGTCATGCGCAGGGCAAGATGTTTAC 533
| | | | |
QY 481 ATCGATTTGCAAGACAGTTGAGGCTTTAATGAAATCATATGTTTCCACTCCCGGAATC 540
| | | | |
DB 534 ATCGATTTGCAAGACAGTTGAGGCTTTAATGAAATCATATGTTTCCACTCCCGGAATC 593
| | | | |
QY 541 TACGACCCGGAAATTCATGCTGCGCCCAACAGAGGAGGCTGCAACAGTGTAAAGGC 600
| | | | |
DB 594 TATGACCCGGAAATTCATGCTGCGCCCAACAGAGGAGGCTGCAACAGTGTAAAGGC 653
| | | | |
QY 601 ACAAGCAAGAGCAAGTTTACGAAATCATCAAGACATCAAGGCTTTAAGAAACCAAC 660
| | | | |
DB 654 ACAAGCAAGAGCAAGTTTACGAAATCATCAAGACATCAAGGCTTTAAGAAACCAAC 713
| | | | |
QY 661 AAAGTGAACAAAGTGTGCTGCTGTGACCTGCAACACAGAGAGTATAGCAATTTGGT 720
| | | | |
DB 714 AAAGTGAACAAAGTGTGCTGCTGTGACCTGCAACACAGAGAGTATAGCAATTTGGT 773
| | | | |
QY 721 GTAGGCTTAATGACACCATGAGAAATCTTTGGCTGCTGAGACAGAAATGAGGCTAG 780
| | | | |
DB 774 GTAGGCTTAATGACACCATGAGAAATCTTTGGCTGCTGAGACAGAAATGAGGCTAG 833
| | | | |
QY 781 ATTTCTCTTCCACCTTGTATGCAATTCCTGTGTGATGAGAAATGTTCTTTCTATAT 840
| | | | |
DB 834 ATTTCTCTTCCACCTTGTATGCAATTCCTGTGTGATGAGAAATGTTCTTTCTATAT 893
| | | | |
QY 841 GGAAGCCTCAGACACATTTTGTACAGAGGCTGATTAATCTTTGCCATGCGAGAAACCT 900
| | | | |
DB 894 GGAAGCCTCAGACACATTTTGTACAGAGGCTGATTAATCTTTGCCATGCGAGAAACCT 953
| | | | |
QY 901 TTGATTTGGAGATGACCTTCAAGAGTGTGACAGACAAATGAATCTGTGTGGTAT 960
| | | | |
DB 954 TTGATTTGGAGATGACCTTCAAGAGTGTGACAGACAAATGAATCTGTGTGGTAT 1013
| | | | |
QY 961 TTTCTTGTGGGGCTGTATCAAGCCAACTCTATAGTTAGTTACAACTCTGGAAAC 1020
| | | | |
DB 1014 TTTCTTGTGGGGCTGTATCAAGCCAACTCTATAGTTAGTTACAACTCTGGAAAC 1073
| | | | |
QY 1021 AATGATGATGAATCTCTGGCTTCAAAACCTTCCGCTCCAGAAATCTCCAAAGGC 1080
| | | | |
DB 1074 AATGATGATGAATCTTCTGGCTTCAAAACCTTCCGCTCCAGAAATCTCCAAAGGC 1133
| | | | |
QY 1081 AACGTTTGGAGATGATGTCACAGCAATGCAATCTCTATAGCTCTGTGACATCCC 1140
| | | | |
DB 1134 AACGTTTGGAGATGATGTCACAGCAATGCAATCTCTATAGCTCTGTGACATCCA 1193
| | | | |
QY 1141 GACCATGTTGTTTATTAATGATGTGCTTAACGTAGGGGATAGCAAGAGCCATGAT 1200
| | | | |
DB 1194 GACCATGTTGTTTATTAATGATGTGCTTAACGTAGGGGATAGCAAGAGCCATGAT 1253
| | | | |
QY 1201 GAGTACACTTCAGAGATTTTCAATGGGTGAAAAGAACACATTTGTTTGCACAAACATGT 1260
| | | | |
DB 1254 GAGTACACTTCAGAGATTTTCAATGGGTGAAAAGAACACATTTGTTTGCACAAACATGT 1313
| | | | |
QY 1261 GAGATTCCTTTTGAAGCTCTTATTAATCTTGAATGATGCTTCTTGTGAGCTGAGC 1320
| | | | |
DB 1314 GAGATTCCTTTTGAAGCTCTTATTAATCTTGAATGATGCTTCTTGTGAGCTGAGC 1373
| | | | |
QY 1321 ACTAGATTCAGTTTAAAGCTGAAAATGAGGGAAAAATTCACATTCACCCAGTTGCT 1380
| | | | |
DB 1374 ACTAGATTCAGTTTAAAGCTGAAAATGAGGGAAAAATTCACATTCACCCAGTTGCT 1433
| | | | |
QY 1381 ACCATTCTCAGCTATCTGACCAAGGCTCCTCTGTTCCACCGGTCACACATGTGGAT 1440
| | | | |
DB 1434 ACCATCTCAGCTATCTGACCAAGGCTCCTCTGTTCCACCGGTCACACATGTGGAT 1493
| | | | |
QY 1441 GCATTCTCAAGAGCGTCAATGCTGAAAACATATAGAGGCTGTGTGGATTGGCC 1500
| | | | |
DB 1494 GCATTCTCAAGAGCGTCAATGCTGAAAACATATAGAGGCTGTGTGGATTGGCC 1553
| | | | |
```

QY 1501 CCAGAGATTAACATGATTCGAGTACAGTGA 1533
DB 1554 CCAGAGATTAACATGATTCGAGTACAGTGA 1586

RESULT 11

US-10-025-003-5
; Sequence 5, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025, 003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-5

Query Match 95.4%; Score 1462.6; DB 14; Length 1533;

Best Local Similarity 97.1%; Pred. No. 0;

Matches 1489; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 1 AGTTCATGAGAAATTTAAGTAGAGAGTCCCTAATGTAAGTACACCGAGCTGAGATT 60
DB 1 AGTTCATGAGAAATTTAAGTAGAGAGTCCCTAATGTAAGTACACCGAGCTGAGATT 60
QY 61 CAGTCCGCTGTAACAATACGAAACCAACGAACTTTGTCAGAGAAACGAAATGACCTAT 120
DB 61 CAGTCCGCTGTAACAATACGAAACCAACGAACTTTGTCAGAGAAACGAAATGACCTAT 120
QY 121 CAGTGAATGTCAAACCAATCCGTCATCAATCAATTTAAACCAACCAACCTATGTTCA 180
DB 121 CAGTGAATGTCAAACCAATCCGTCATCAATCAATTTAAACCAACCAACCTATGTTCT 180
QY 181 AATTTGGGGGTATGTTTGGGTGGGTGGGTAACAAACGGCTTACCTCAACGGTGGT 240
DB 181 AATTTGGGGGTATGTTTGGGTGGGTGGGTAACAAACGGCTTACCTCAACGGTGGT 240
QY 241 GTTATGCTTAACAGAGAGACATTTTCATGGGCTTAAAGAGACAAAGATTCAAACGCAAT 300
DB 241 GTTATGCTTAACAGAGAGACATTTTCATGGGCTTAAAGAGACAAAGATTCAAACGCAAT 300
QY 301 TACTTTGGCTCCCTCAACCCAGGCTCAGCTATCCAGTTCGAGTTGGGTCTTCCAGGGAAGAA 360
DB 301 TACTTTGGCTCCCTCAACCCAGGCTCAGCTATCCAGTTCGAGTTGGGTCTTCCAGGGAAGAA 360
QY 361 ATCTATGCCCCATTAAGAGTCTGCTTCCAAATGTTTAACTCTGACACATTTGTTGGG 420
DB 361 ATCTATGCCCCATTAAGAGTCTGCTTCCAAATGTTTAACTCTGACACATTTGTTGGG 420
QY 421 GGATGGATATGACAACTGAACCTGGCTGATGCGATGGGCGAGGCAAAAGGTTTGAC 480
DB 421 GGATGGATATGACAACTGAACCTGGCTGATGCGATGGGCGAGGCAAAAGGTTTGAC 480
QY 481 ATCGATTTGACAAAGAGCTTGAAGCTTTACATGGAATTCATGTTCCACTCCCGGAATC 540
DB 481 ATCGATTTGACAAAGAGCTTGAAGCTTTACATGGAATTCATGTTCCACTCCCGGAATC 540
QY 541 TACGACCCGGATTTTCACTTGTCTGCAACCAAGAGAGCGTCCCAACAGTGAATGAAGGC 600

DB 541 TATGACCCGGATTTTCACTTGTCTGCAACCAAGAGAGCGTCCCAACAGTGAATGAAGGC 600
QY 601 ACAAAGCAAGAGCAAGTTTCAAGAAATCATCAAGACATCAAGAGCTTTAAGAAAGCAAC 660
DB 601 ACAAAGCAAGAGCAAGTTTCAAGAAATCATCAAGACATCAAGAGCTTTAAGAAAGCAAC 660
QY 661 AAGTGAACAAAGTGTGTCTGTGAGCTGCAACAGAGAGTATGCAATTTGGT 720
DB 661 AAGTGAACAAAGTGTGTGTGTGAGCTGCAACAGAGAGTATGCAATTTGGT 720
QY 721 GTAGGCTTAATGACACCATGAGAAATCTCTGGCTGCTGTGAGACAGAAATGAGCTGAG 780
DB 721 GTAGGCTTAATGACACCATGAGAAATCTCTGGCTGCTGTGAGACAGAAATGAGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGTATGCCATTCCTGTGTGATGAAATGTTCTTTCAATAT 840
DB 781 ATTTCTCTTCCACCTTGTATGCCATTCCTGTGTGATGAAATGTTCTTTCAATAT 840
QY 841 GGAAGCCCTCAGAACACCTTTGTACAGAGGCTGATGATCTTGGCATCGCGAGAACAT 900
DB 841 GGAAGCCCTCAGAACACCTTTGTACAGAGGCTGATGATCTTGGCATCGCGAGAACAT 900
QY 901 TTGATTTGGTGAATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGAT 960
DB 901 TTGATTTGGTGAATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGAT 960
QY 961 TTTCTTGTGGGGGTGTATCAAGCCAAATCTATAGTTAGTTACAAACCATCTGGGAAC 1020
DB 961 TTTCTTGTGGGGGTGTATCAAGCCAAATCTATAGTTAGTTACAAACCATCTGGGAAC 1020
QY 1021 AATGATGTAATATCTCTGGCTTCAACAACTTCCGCTCCAGAAATCTTCAAGAGC 1080
DB 1021 AATGATGTAATATCTCTGGCTTCAACAACTTCCGCTCCAGAAATCTTCAAGAGC 1080
QY 1081 AAGTTTGTGAATATGATGTCACAGCAATGTCATCTTATAGAGCTGTGTAACATCCA 1140
DB 1081 AAGTTTGTGAATATGATGTCACAGCAATGTCATCTTATAGAGCTGTGTAACATCCA 1140
QY 1141 GACCAATGTTGTTATTAAGTATGTCCTTACGTAGGGATGACAAAGAGCCATGAT 1200
DB 1141 GACCAATGTTGTTATTAAGTATGTCCTTACGTAGGGATGACAAAGAGCCATGAT 1200
QY 1201 GAGTACACTTCAAGATATTCATGCGGTGAAAGAACACCATTTGTTGCAACAATGT 1260
DB 1201 GAGTACACTTCAAGATATTCATGCGGTGAAAGAACACCATTTGTTGCAACAATGT 1260
QY 1261 GAGATTCCTCTTTAGCTGCTCTATTAATCTTGAATTTGGAATTTCTTGTGAGCTCAGC 1320
DB 1261 GAGATTCCTCTTTAGCTGCTCTATTAATCTTGAATTTGGAATTTCTTGTGAGCTCAGC 1320
QY 1321 ACTAAGATTCAGTTTAAAGCTGAAATGAGGAAATTCATCTCAATCCACAGTGTCT 1380
DB 1321 ACTAAGATTCAGTTTAAAGCTGAAATGAGGAAATTCATCTCAATCCACAGTGTCT 1380
QY 1381 ACCATTCCTCACTATCTGACCAAGGCTCTCTGTTTCCACCGGATACACAGTGTGAT 1440
DB 1381 ACCATTCCTCACTATCTGACCAAGGCTCTCTGTTTCCACCGGATACACAGTGTGAT 1440
QY 1441 GATTTGTCAAAAGCAGCGGCAATGCTGGAATAAATTAAGAGGCTTTGTTGATTTGCC 1500
DB 1441 GATTTGTCAAAAGCAGCGGCAATGCTGGAATAAATTAAGAGGCTTTGTTGATTTGCC 1500
QY 1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533
DB 1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533

RESULT 12

US-10-718-952-5
; Sequence 5, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:

/ APPLICANT: Hitz, William
 / APPLICANT: Sebastian, Scott
 / APPLICANT: Grace, John
 / APPLICANT: Streitz, Leon
 / TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
 / TITLE OF INVENTION: SACCHARIDES AND PHYTIC ACID
 / FILE REFERENCE: BR-1077-C
 / CURRENT APPLICATION NUMBER: US/10/718,952
 / CURRENT FILING DATE: 2003-11-21
 / PRIOR APPLICATION NUMBER: 08/835,751
 / PRIOR FILING DATE: APRIL 8, 1997
 / PRIOR APPLICATION NUMBER: PCT/US98/06822
 / PRIOR FILING DATE: APRIL 7, 1998
 / NUMBER OF SEQ ID NOS: 16
 / SOFTWARE: Microsoft Office 97
 / SEQ ID NO 5
 / LENGTH: 1533
 / TYPE: DNA
 / ORGANISM: Glycine max
 / US-10-718-952-5

Query Match	95.4%	Score 1462.6;	DB 19;	Length 1533;
Best Local Similarity	97.1%;	Pred. No. 0;		
Matches 1489; Conservative	0;	Mismatches 44;	Indels 0;	Gaps 0;

Qy	1	ATGTCATCGAAGATTTTAAAGGTAGAGATCCTTAATGGAAGTACACCGAATCGAAT	60
Db	1	ATGTTCAATCGAAGATTTTAAAGGTAGAGTCTTAATGGAAGTACACCGAATCGAAT	60
Qy	61	CAGTCGGTGTACACTAGAAACCAACCGAATCTTGTTCAAGAAACAGGAATGGACCTAT	120
Db	61	CAGTCGGTGTACAACACTAGAAACCAACCGAATCTTGTTCAAGAAACAGGAATGGACCTAT	120
Qy	121	CAGTGAATTGTCAAAACCCAAATCTGCAAAATGCAATTTAAACCAACACCATGTTCCA	180
Db	121	CAGTGAATTGTCAAAACCCAAATCTGCAAAATGCAATTTAAACCAACATCATGTTCTT	180
Qy	181	AAATTTGGGGGTGATGCTTGTTGGGTTTGGGGTGGAAAACAACGGCTCTAACCCTCACGGTGT	240
Db	181	AAATTTAGGGGTATATGCTTGTTGGGTTTGGGGTGGAAAACAACGGCTCAACCTTCACGGTGT	240
Qy	241	GTTATTTGCTPACAGAGAGACATTTTCATGGGCTTAAGAAAGCAAGATTCAACAGCCAAT	300
Db	241	GTTATTTGCTPACAGAGAGGCAATTTTCATGGGCTTAAGAAAGCAAGATTCAACAGCCAAT	300
Qy	301	TACTTTGGCTCCCTCAACCCAGGCTCAGCTATTTCAGTTGATGATCTTCCAGAGGAGAGAA	360
Db	301	TACTTTGGCTCCCTCAACCCAGGCTCAGCTATCCGATTTGGGTTCTTCAAGGAGAGAGAA	360
Qy	361	ATCTATGCCCATTTCAAGAGTCTGCTTCCATATGTTAATCCTGACAGATTTGTGTTGGG	420
Db	361	ATCTATGCCCATTTCAAGAGCTGCTTCCATATGTTAATCCCTGACAGATTTGTGTTGGG	420
Qy	421	GGATGGGATATCAGCAACATGAACTGGGTGATGTCATGSCCAAGGGCAAAAGTGTGGAC	480
Db	421	GGATGGGATATCAGCAACATGAACTGGGTGATGTCATGSCCAAGGGCAAAAGTGTGGAC	480
Qy	481	ATCGATTTGCAGAGCAGTTTGAAGCCTTAATCATGAAATTCATATGTTCTACCTCCCGGAATC	540
Db	481	ATCGATTTGCAGAGCAGTTTGAAGCCTTTCATGAAATTCATATGCTTCACTCCCGGAATC	540
Qy	541	TACGACCCCGGATTTCAATTTGTGTCGCAACGAAGAGAGCGTGCACAAAGCTATTAAGGGC	600
Db	541	TATGACCCCGGATTTCAATTTGTGTCGCAACGAAGAGAGCGTGCACAAAGCTATTAAGGGC	600
Qy	601	ACAAAGCAGAGCAGATTTCAAGAAATCATCAAGAATCAAGCGCTTTTAAAGAACCAAC	660
Db	601	ACAAAGCAGAGCAGATTTCAACAAATCATCAAGAATCAAGCGCTTTTAAAGAACCAAC	660
Qy	661	AAAGTGCACAAAGTGTGTCTCTGTGGAATCTGCCAACACAGAGAGGTATAGCAATTTGGTT	720
Db	661	AAAGTGCACAAAGTGTGTCTCTGTGGAATCTGCCAACACAGAGAGGTATAGTAATTTGGTT	720

QY	721	GTAGGCGCTTAATGACACCAAGGAGAAATCTCTTGAGCTGCTGTGACAGAAATGAGGCTGAG	780
Db	721	GTGGGCGCTTAATGACACCAATGGAATCTCTTGAGCTGCTGTGACAGAAATGAGGCTGAG	780
QY	781	ATTTCCTCTTCACACTTGTGATGCCATGCTGTGTGATGGAATAATGTCTCTTCATTAAAT	840
Db	781	ATTTCCTCTTCACACTTGTGATGCCATGCTGTGTGATGGAATAATGTCTCTTCATTAAAT	840
QY	841	GGAAGCGCTCAGAACACTTTTGTATCCAGGCGTGTATCTTGCGCATCCGAGGAACT	900
Db	841	GGAAGCGCTCAGAACACTTTTGTATCCAGGCGTGTATCTTGCGCATCCGAGGAACT	900
QY	901	TTGATGTGTGAGATGACTTCAAGAGTGTCAAGCCAAATGAAATCTGTGTGTTGAT	960
Db	901	TTGATGTGTGAGATGACTTCAAGAGTGTCAAGCCAAATGAAATCTGTGTGTTGAT	960
QY	961	TTTCTCTGTGGGGCGCTGGTATCCAGGCCAACATCTATAGTTAGTTACCAACCATCTGGGAAAC	1020
Db	961	TTTCTCTGTGGGGCGCTGGTATCCAGGCCAACATCTATAGTTAGTTACCAACCATCTGGGAAAC	1020
QY	1021	AATGATGGTATGAAATCTCTCGGCGTCCAAACACTTCGCGTCCAGAGAAATCTCCAAGAGC	1080
Db	1021	AATGATGGTATGAAATCTCTCGGCGTCCAAACACTTCGCGTCCAGAGAAATCTCCAAGAGC	1080
QY	1081	AACGTTGTGACGATATGTCTACAGCAATGCCATCCTCTATGAGCCTGTGTGAACATCCC	1140
Db	1081	AACGTTGTGATGATATGTCTACAGCAATGCCATCCTCTATGAGCCTGTGTGAACATCCA	1140
QY	1141	GACCATTGTTGTTATTTAAAGTATGTGCTTACGTAGGGGATGACAGAAAGCATGTGAT	1200
Db	1141	GACCATTGTTGTTATTTAAAGTATGTGCTTACGTAGGGGATGACAGAAATGAGCATGTGAT	1200
QY	1201	GAGTACACTTCAGAGATATTCAATGGGTGGAAGAAACACATGTTTGTGACAAACATCATGT	1260
Db	1201	GAGTACACTTCAGAGATATTCAATGGGTGGAAGAAAGCACATGTTTGTGACAAACATCATGC	1260
QY	1261	GAGGATTCCTCTTTAGCTGCTCCTAATTATCTTGACTTGTGCTTCTTCTGCTGAGCTGAGC	1320
Db	1261	GAGGATTCCTCTTTAGCTGCTCCTAATTATCTTGACTTGTGCTTCTTCTGCTGAGCTGAGC	1320
QY	1321	ACTGTAATCCAGTTTAAGCTGAAAAAGAGGAAAAATTCACATCATTCACCCAGTTGCT	1380
Db	1321	ACTGTAATCCAGTTTAAGCTGAAAAAGAGGAAAAATTCACATCATTCACCCAGTTGCT	1380
QY	1381	ACCAATCTCAGCTATCTGACCAAGGCTCTCTGTTTCAACCGGGTACACCAAGTGTGAAT	1440
Db	1381	ACCAATCTCAGCTATCTGACCAAGGCTCTCTGTTTCAACCGGGTACACCAAGTGTGAAT	1440
QY	1441	GCATTGTCAAAAGCAGCGTGCATCTCTGAAAAACATAATGAGGCGCTGTGTGTGATTGGCC	1500
Db	1441	GCATTGTCAAAAGCAGCGTGCATCTCTGAAAAACATAATGAGGCGCTGTGTGTGATTGGCC	1500
QY	1501	CCAGAGATAACATGATTTCTGAGTAAACATGTA	1533
Db	1501	CCAGAGATAACATGATTTCTGAGTAAACATGTA	1533

```

RESULT 13
US-10-424-599-70167
; Sequence 70167, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 70167

```



```

; LENGTH: 1989
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1
US-10-424-599-70167

Query Match      95.0%; Score 1456.4; DB 18; Length 1989;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1492; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 1 ATGTCATCGAGAAATTTTAAGTAGAGAGTCTAATGTGAATACACCGAGCTGAGATT 60
DB 217 ATGTCATCGAGAAATTTTAAGTAGAGTCTAATGTGAATACACCGAGCTGAGATT 276
QY 61 CAGTCCGCTGACAACTACGAAACCAACGAACTTGTTCACGAGAACAGAAATGGCACTAT 120
DB 277 CAGTCCGCTGACAACTACGAAACCAACGAACTTGTTCACGAGAACAGAAATGGCACTAT 336
QY 121 CAGTGAATGTCAAAACCCAAATCCGTCACTACATTTTAAACCAACCCATGTTCCA 180
DB 337 CAGTGAATGTCAAAACCCAAATCTGTCAATACGAAATTTAAACCAACATCCATGTTCC 396
QY 181 AATTTGGGGGTGATGCTTGTGGGT- GGGGTGGAACAACGCTCTACCTCACCGGTG 239
DB 397 AATTTGGGGGTATGCTTGTGGGTGGGGTGAACAACGCTCTACCTCACCGGTG 456
QY 240 TGTATTGCTAACAGAGAGACATTTTCATGGGCTACAAAGACAGATTTCAACAGCCAA 299
DB 457 TGTATTGCTAACAGAGAGGCAATTTTCATGGGCTACAAAGACAGATTTCAACAGCCAA 516
QY 300 TTAATTTGGCTCCCTACCAAGCTCAGTATTCAGATTGATCCTTCCAGGAGAGGA 359
DB 517 TTAATTTGGCTCCCTACCAAGCTCAGTATTCAGATTGATCCTTCCAGGAGAGGA 576
QY 360 AATCTATGCCCATTCAGAGATGCTCTCAATGTGTTAATCCGACGACATTTGTGTTG 419
DB 577 AATCTATGCCCATTCAGAGATGCTCTCAATGTGTTAATCCGACGACATTTGTGTTG 636
QY 420 GGGATGGATATCAGAACATGAACCTGCTGATGTCATGAGCCAGGCAAAAGTGTGTA 479
DB 637 GGGATGGATATCAGAACATGAACCTGCTGATGTCATGAGCCAGGCAAAAGTGTGTA 696
QY 480 CATCGATTTGCAAGACAGTGAAGCCTTACATGATTCATGTTCCATCCCGGAAT 539
DB 697 CATCGATTTGCAAGACAGTGAAGCCTTACATGATTCATGTTCCATCCCGGAAT 756
QY 540 CTACGACCCGGAATTTCAATGCTGCAACCAAGAGAGCGGTGCCAACAGGTGTTAAG 599
DB 757 CTATGACCCGGAATTTCAATGCTGCAACCAAGAGAGCGGTGCCAACAGGTGTTAAG 816
QY 600 CACAAAGCAAGAGCAAGTTCAGCAATTCATCAAGACATCAAGCGCTTTAAGAACCCAC 659
DB 817 CACAAAGCAAGAGCAAGTTCAGCAATTCATCAAGACATCAAGCGCTTTAAGAACCCAC 876
QY 660 CAAAGTGAACAAGTGTGTCTCTGTGACCTGCAACAAGAGAGTATAGCAATTTGGT 719
DB 877 CAAAGTGAACAAGTGTGTCTCTGTGACCTGCAACAAGAGAGTATAGCAATTTGGT 936
QY 720 TGTAGGCTTTAATGACACCATGAGAAATCTTGGCTGCTGAGAGAAATGAGGCTGA 779
DB 937 TGTAGGCTTTAATGACACCATGAGAAATCTTGGCTGCTGAGAGAAATGAGGCTGA 996
QY 780 GATTTCTCTTCACCTTGATGCAATTCCTGTGTGATGAGAAATGTTCTTTCAATTA 839
DB 997 GATTTCTCTTCACCTTGATGCAATTCCTGTGTGATGAGAAATGTTCTTTCAATTA 1056
QY 840 TGAAGACCTTCAGAACATTTTGTATCAGAGGCTGATTTGATCTTGCAATGCCAGAACAC 899
DB 1057 TGAAGACCTTCAGAACATTTTGTATCAGAGGCTGATTTGATCTTGCAATGCCAGAACAC 1116
QY 900 TTGATGTGTGAGATGATCTCAAGAGGTGTCAGCAAAATGAATCTGTGTGTTGA 959
DB 1117 TTGATGTGTGAGATGATCTCAAGAGGTGTCAGCAAAATGAATCTGTGTGTTGA 1176
QY 960 TTTTCTTTGGGGGCTGGTATCAAGCCAACTATATGTTAGTTACACCATCTGGGAAA 1019
DB 1177 TTTTCTTTGGGGGCTGGTATCAAGCCAACTATATGTTAGTTACACCATCTGGGAAA 1236
QY 1020 CAATGATGATGAATCTCTGGGCTCAACAACTTCGCTCCAGAGAAATCTCCAAAG 1079
DB 1237 CAATGATGATGAATCTCTGGGCTCAACAACTTCGCTCCAGAGAAATCTCCAAAG 1296
QY 1080 CAACGTTGTTGACGATATGATGATCAACAGCAATGCAATCTCTATGAGCTGTGAACATCC 1139
DB 1297 CAACGTTGTTGATGATATGATGATCAACAGCAATGCAATCTCTATGAGCTGTGAACATCC 1356
QY 1140 CGACCATGTTGTTGTTATTAATGATGCTTACATGAGGGATAGCAAGAGCCATGGA 1199
DB 1357 AGACCATGTTGTTGTTATTAATGATGCTTACATGAGGGATAGCAAGAGCCATGGA 1416
QY 1200 TGAGTACATTCAGAGATATTCATGAGGTGAAAAGAACACATGTTTTCACAAACATG 1259
DB 1417 TGAGTACATTCAGAGATATTCATGAGGTGAAAAGAACACATGTTTTCACAAACATG 1476
QY 1260 TGAGATTCCTTTTATGCTGCTCTATTATCTTGAATCTGTCCTTGTGAGCTGAG 1319
DB 1477 CGAGATTCCTTTTATGCTGCTCTATTATCTTGAATCTGTCCTTGTGAGCTGAG 1536
QY 1320 CACTAGATTCAGTTTAAAGCTGAAAATGAGGAAAATTCACCTCATTCACCCAGTTGC 1379
DB 1537 CACTAGATTCAGTTTAAAGCTGAAAATGAGGAAAATTCACCTCATTCACCCAGTTGC 1596
QY 1380 TACCAATTCACATCTACGACCAAGGCTCTGCTGCTCAACGGGTGACACAGTGTGAA 1439
DB 1597 TACCAATTCACATCTACGACCAAGGCTCTGCTGCTCAACGGGTGACACAGTGTGAA 1656
QY 1440 TGCATTTGCAAGACGCTGCAATGCTGGAACCAATATGAGGCTTGTGTGATTTG 1499
DB 1657 TGCATTTGCAAGACGCTGCAATGCTGGAACCAATATGAGGCTTGTGTGATTTG 1716
QY 1500 CCCAGAGAAATTAATGATTTCTGAGTACAGTGA 1533
DB 1717 CCCAGAGAAATTAATGATTTCTGAGTACAGTGA 1750
```

```

DB 1117 TTGATGTGTGAGATGATCTCAAGAGGTGTCAGCAAAATGAATCTGTGTGTTGA 1176
QY 960 TTTTCTTTGGGGGCTGGTATCAAGCCAACTATATGTTAGTTACACCATCTGGGAAA 1019
DB 1177 TTTTCTTTGGGGGCTGGTATCAAGCCAACTATATGTTAGTTACACCATCTGGGAAA 1236
QY 1020 CAATGATGATGAATCTCTGGGCTCAACAACTTCGCTCCAGAGAAATCTCCAAAG 1079
DB 1237 CAATGATGATGAATCTCTGGGCTCAACAACTTCGCTCCAGAGAAATCTCCAAAG 1296
QY 1080 CAACGTTGTTGACGATATGATGATCAACAGCAATGCAATCTCTATGAGCTGTGAACATCC 1139
DB 1297 CAACGTTGTTGATGATATGATGATCAACAGCAATGCAATCTCTATGAGCTGTGAACATCC 1356
QY 1140 CGACCATGTTGTTGTTATTAATGATGCTTACATGAGGGATAGCAAGAGCCATGGA 1199
DB 1357 AGACCATGTTGTTGTTATTAATGATGCTTACATGAGGGATAGCAAGAGCCATGGA 1416
QY 1200 TGAGTACATTCAGAGATATTCATGAGGTGAAAAGAACACATGTTTTCACAAACATG 1259
DB 1417 TGAGTACATTCAGAGATATTCATGAGGTGAAAAGAACACATGTTTTCACAAACATG 1476
QY 1260 TGAGATTCCTTTTATGCTGCTCTATTATCTTGAATCTGTCCTTGTGAGCTGAG 1319
DB 1477 CGAGATTCCTTTTATGCTGCTCTATTATCTTGAATCTGTCCTTGTGAGCTGAG 1536
QY 1320 CACTAGATTCAGTTTAAAGCTGAAAATGAGGAAAATTCACCTCATTCACCCAGTTGC 1379
DB 1537 CACTAGATTCAGTTTAAAGCTGAAAATGAGGAAAATTCACCTCATTCACCCAGTTGC 1596
QY 1380 TACCAATTCACATCTACGACCAAGGCTCTGCTGCTCAACGGGTGACACAGTGTGAA 1439
DB 1597 TACCAATTCACATCTACGACCAAGGCTCTGCTGCTCAACGGGTGACACAGTGTGAA 1656
QY 1440 TGCATTTGCAAGACGCTGCAATGCTGGAACCAATATGAGGCTTGTGTGATTTG 1499
DB 1657 TGCATTTGCAAGACGCTGCAATGCTGGAACCAATATGAGGCTTGTGTGATTTG 1716
QY 1500 CCCAGAGAAATTAATGATTTCTGAGTACAGTGA 1533
DB 1717 CCCAGAGAAATTAATGATTTCTGAGTACAGTGA 1750

RESULT 14
US-10-424-599-12022
; Sequence 12022, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12022
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110863C.1
US-10-424-599-12022

Query Match      82.3%; Score 1261; DB 18; Length 2018;
Best Local Similarity 88.9%; Pred. No. 0;
Matches 1363; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 1 ATGTCATCGAGAAATTTTAAGTAGAGAGTCTAATGTGAATACACCGAGCTGAGATT 60
DB 98 ATGTCATCGAGAGATTTCAGAGTTGAGAGTCTTAACGTGAAGTACACAGAGCTAGATT 157
```

```
QY 61 CAGTCCGTGTAACTACGAAACCAACGAACTTGTTCAGAGAAACAGATGGACCTAT 120
DB 158 CAGTCCGTGTAACTACGAAACCAACGAACTTGTTCAGAGAAACAGATGGACCTAT 217
QY 121 CAGTGGATTTGCAAAACCAAAATCCGTCAACATCAATTTAAACCAACCAATGTTCCA 180
DB 218 CAGTGGATTTGCAAAACCAAAATCCGTCAACATCAATTTAAACCAACCAATGTTCCA 277
QY 181 AAATTTGGGGGTGATGTTTGGGTTGGGGTGAANAAGGGCTTACCTTCACCGGTGAT 240
DB 278 AAATTTGGGGGTGATGTTTGGGTTGGGGTGAANAAGGGCTTACCTTCACCGGTGAT 337
QY 241 GTTATTTGCTAAGAGAGACATTTTCATGGGCTACAAAGACAGATTTCAACAGCAAT 300
DB 338 GTTATTTGCTAAGAGAGAGATTTTCATGGGCTACAAAGACAGATTTCAACAGCAAT 397
QY 301 TACTTTGGCTCCCTCAACCCAGGCTCAGCTTATTCAGATTGATCTTTCAGAGAGAGAA 360
DB 398 TACTTTGGCTCCCTCAACCCAGGCTCAGCTTATTCAGATTGATGGGGTCTTTCAGAGAGAA 457
QY 361 ATCTAAGCCCAATCAAGAGCTGCTTCCAAATGCTTAATCTGACGACATTTGTTGGG 420
DB 458 ATATATGCTCCATCAAGAGCTGCTTCCAAATGCTTAATCTGATGATGTTGTTGGG 517
QY 421 GGATGGGATATCAGCAACATGAACCTGGCTGATGCGATGGCCAGAGGCAAGGTTTGAC 480
DB 518 GGATGGGATATCAGTAACCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 577
QY 481 ATCGATTTGCAAGAGAGTTGAGGCTTACATGAAATCCATGCTTCCCTCCGGAATC 540
DB 578 ATCGAATCTCAGAAACAGTTGAGGCGGTACATGAAATCCATGCTTCCCTCCGGAATC 637
QY 541 TACGACCCGAGTTTCTTCTGCTGCAACCAAGAGAGGCTGCAACCAAGTATTAAGGCG 600
DB 638 TATGACCCGAGTTTCTTCTGCTGCAACCAAGAGAGGCTGCAACCAAGTATTAAGGCG 697
QY 601 ACNAAAGCAGAGAGAGTTGAGCAAAATCATCAAAAGACATCAAGGCTTTAAGAGAGCACC 660
DB 698 ACNAAAGAGAGAGAGTTGAGCAAAATCATCAAAAGACATTAAGGAGTTCAAGAGAGACCT 757
QY 661 AAAGTGACAGAGGTTGTTCTGCTGAGCTGCAACAGAGAGGATTAAGCAATTTGGTT 720
DB 758 AAAGTTGACAGAGGTTGTTCTGCTGAGCTGCAACAGAGAGGATTAAGCAATTTGGTT 817
QY 721 GTAGGCTTAATGACCAATGAGGATCTTGGCTGCTGAGACAGAAATGAGGCTGAG 780
DB 818 GTAGGCTTAATGACCAATGAGGATCTTGGCTGCTGAGACAGAAATGAGGCTGAG 877
QY 781 ATTTCTCTTCCACCTTGTATGCTGCTGCTGATGGAATAATGTTCTTTTCAATTAAT 840
DB 878 ATTTCTCTTCCACCTTGTATGCTGCTGCTGATGGAATAATGTTCTTTTCAATTAAT 937
QY 841 GGAAGCCTCAGAAACCTTTTGTACAGAGGCTGATGATCTTGCATGCGAGAGAACTAT 900
DB 938 GGAAGCCTCAGAAACCTTTTGTACAGAGCTATGATCTTGCATGGAATGGAATGAT 997
QY 901 TTGATTTGGGAGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTTGGTTGAT 960
DB 998 TTGATTTGGGAGATGATCTTCAAGAGTGTGAGACCAAAATGAAATCTGTTGGTTGAT 1057
QY 961 TTTCTTGGGGGCTGTATCAAGCAACATCTATAGTTAGTTAACAACCTTGGGAAAC 1020
DB 1058 TTTCTTGGGGGCTGTATCAAGCAACATCTATAGTTAGTTAACAACCTTGGGAAAC 1117
QY 1021 AATGATGATGATGATCTTGGGCTCAGAAACCTTCCGCTCAAGGAAATCTTCCAGAGC 1080
DB 1118 AATGATGATGATGATCTTGGGCTCAGAAACCTTCCGCTCAAGGAAATCTTCCAGAGC 1177
QY 1081 AAGTTGTTGAGATATGTTCAACAGCAATGCTATCTATAGAGCTGTTGAACATCCC 1140
DB 1178 AAGTTGTTGAGATATGTTCAACAGCAATGCTATCTATAGAGCTGTTGAACATCCC 1237
```

```
QY 1141 GACCATGTTGTTGTTATTAAGTATGCTTACGTAAGGGGATTAAGCAAGAGACCATGAT 1200
DB 1238 GACCATGTTGTTGTTATTAAGTATGCTTACGTAAGGGGATTAAGCAAGAGACCATGAT 1297
QY 1201 GAGTACATCTTCAAGATATTTCAATGGGTTGAAAGAACACATTTGTTTTCAGAAACATG 1260
DB 1298 GAGTACATCTTCAAGATATTTCAATGGGTTGAAAGAACACATTTGTTTTCAGAAACATG 1357
QY 1261 GAGATTCCTTTTGTGCTGCTCTTATTAATCTTGAATGCTGCTTCTTCTGAGCTGAGC 1320
DB 1358 GAGATTCCTTGTGCTGCTCTTATTAATCTTGAATGCTGCTTCTTCTGAGCTGAGC 1417
QY 1321 ACTAGAAATCCAGTTTAAAGCTGAAAGAGGAAATTTCACTCATTCACCATGCTGCT 1380
DB 1418 ACTAGAAATCCAGTTTAAAGCTGAAAGAGGAAATTTCACTCATTCACCATGCTGCTG 1477
QY 1381 ACCATTCCTGATATCTGATCCAGAGCTCTCTGTTTCCACCGGGATCAACAGTGTGAT 1440
DB 1478 ACCATTCCTGATATCTGATCCAGAGCTCTCTGTTTCCACCGGGATCAACAGTGTGAT 1537
QY 1441 GCATTTGCAAGCAGCTGCAATGCTGGAATAACATTAATGAGGCTTGTGTTGATTTGCC 1500
DB 1538 GCATTTGCAAGCAGCTGCAATGCTGGAATAACATTAATGAGGCTTGTGTTGATTTAGCT 1597
QY 1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533
DB 1598 CCTGAGAAACATGATCTCTGAGTACAAAGTGA 1630
```

```
RESULT 15
US-10-424-599-12021/c
; Sequence 12021, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424.599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12021
; LENGTH: 2582
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110862C.1
US-10-424-599-12021
```

```
Query Match 76.7%; Score 1176.4; DB 18; Length 2582;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1183; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ANGTTCATCGAGAAATTTTAAGGTAGAGAGTCTTAATGTAAGTACACGAGCTGAGATT 60
DB 2454 ANGTTCATCGAGAAATTTTAAGGTAGAGAGTCTTAATGTAAGTACACGAGCTGAGATT 2395
QY 61 CAGTCCGTGTAACTACGAAACCAACGAACTTGTTCAGAGAAACAGATGGACCTAT 120
DB 2394 CAGTCCGTGTAACTACGAAACCAACGAACTTGTTCAGAGAAACAGATGGACCTAT 2335
QY 121 CAGTGGATTTGCAAAACCAAAATCCGTCAACATCAATTTAAACCAACCAATGTTCCA 180
DB 2334 CAGTGGATTTGCAAAACCAAAATCCGTCAACATCAATTTAAACCAACCAATGTTCCA 2275
QY 181 AAATTTGGGGGTGATGTTTGGGTTGGGGTGAANAAGGGCTTACCTTCACCGGTGAT 240
DB 2274 AAATTTGGGGGTGATGTTTGGGTTGGGGTGAANAAGGGCTTACCTTCACCGGTGAT 2215
QY 241 GTTATTTGCTAAGAGAGATTTTCATGGGCTACAAAGAGACATTAAGCAAT 300
```

Db 2214 GTTATTGCTAAGAGAGGCAATTTCTATGGGCTACAAAGSACAAGATTCACAAGCCAAAT 2155
OY 301 TACTTTGGCTCCCTCAACCCCAAGCCTCACTATTGAGTTGGATCTTCCAGGGAGAGAA 360
Db 2154 TACTTTGGCTCCCTCAACCCCAAGCCTCACTATTGAGTTGGATCTTCCAGGGAGAGAA 2095
OY 361 ATCTATGCCCATTCAGAGAGTGTGCTTCCAAATGGTTAAATCCAGAGCATATGTGTTGGG 420
Db 2094 ATCTATGCCCATTCAGAGAGTGTGCTTCCAAATGGTTAAATCCAGAGCATATGTGTTGGG 2035
OY 421 GGATGGGATATCAGACAATGAACCTGCTGATGCCATGGCCAGGGCAAAAGTGTGGAC 480
Db 2034 GGATGGGATATCAGACAATGAACCTGCTGATGCCATGGCCAGGGCAAAAGTGTGGAC 1975
OY 481 ATCGATTTCAGAGAGAGTTCAGAGGCTTACATGAAATCCATGGTTCCACTCCCGGAATC 540
Db 1974 ATCGATTTCAGAGAGAGTTCAGAGGCTTACATGAAATCCATGGTTCCACTCCCGGAATC 1915
OY 541 TACGACCCGATTTCTATGCTGCAACCAAGAGAGCGTCCCAACACGTGATTAAGGCG 600
Db 1914 TACGACCCGATTTCTATGCTGCAACCAAGAGAGCGTCCCAACACGTGATTAAGGCG 1855
OY 601 ACAAGACAAGAGCAAGTTCAGCAATCATCAAGAATCAAGSGCTTTAAGAAAGCAAC 660
Db 1854 ACAAGACAAGAGCAAGTTCAGCAATCATCAAGAATCAAGSGCTTTAAGAAAGCAAC 1795
OY 661 AAAGTGACAGAGTGTGCTGCTGAGACTGCAACACAGAGAGTATAGCAATTTGGTT 720
Db 1794 AAAGTGACAGAGTGTGCTGCTGAGACTGCAACACAGAGAGTATAGCAATTTGGTT 1735
OY 721 GTAAGGCTTAAATGACACATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
Db 1734 GTAAGGCTTAAATGACACATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 1675
OY 781 ATTTCCCTTCCACCTTGTATGCAATTCCTGCTGTGATGGAAATGTTCCCTTCATTAAAT 840
Db 1674 ATTTCCCTTCCACCTTGTATGCAATTCCTGCTGTGATGGAAATGTTCCCTTCATTAAAT 1615
OY 841 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTCCTTGCATTCGAGAGAACAT 900
Db 1614 GGAAGCCCTCAGAACCTTTTGTACAGAGGCTGATTCCTTGCATTCGAGAGAACAT 1555
OY 901 TTGATTGGTGAATGATCTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGGTTGAT 960
Db 1554 TTGATTGGTGAATGATCTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGGTTGAT 1495
OY 961 TTTCTGTGGGGCTGTGTCAAGCAACATCTATAGTTAGTACAAACCATCTGGGAAAC 1020
Db 1494 TTTCTGTGGGGCTGTGTCAAGCAACATCTATAGTTAGTACAAACCATCTGGGAAAC 1435
OY 1021 AATGATGTATGATCTCTCGGCTCCAAACCTTCGCTCCAAAGGAAATCTTCCAGAGC 1080
Db 1434 AATGATGTATGATCTCTCGGCTCCAAACCTTCGCTCCAAAGGAAATCTTCCAGAGC 1375
OY 1081 AAGCTTTTGAAGATATGTGTCAACAGCAATGCAATCTCTATAGAGCTGTGAAATCCC 1140
Db 1374 AAGCTTTTGAAGATATGTGTCAACAGCAATGCAATCTCTATAGAGCTGTGAAATCCA 1315
OY 1141 GACCATGTTGTTGTTATTAAGTATGACCTTACGTAGGGGATGCAAGAGAGCC 1194
Db 1314 GACCATGTTGTTGTTATTAAGTATGACCTTACGTAGGGGATGCAAGAGAGCC 1261

Search completed: June 8, 2005, 11:26:00
Job time : 958.946 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model1

Run on: June 7, 2005, 21:02:47 ; Search time 267.095 Seconds
(without alignments)
9570.620 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533

Sequence: 1 agctcaccgagacattta.....tgatctcgagacaaagta 1533

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/5B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	984.2	64.2	1959	4	US-09-727-628-1
2	979.4	63.9	1931	3	US-09-118-442-10
3	979.4	63.9	1931	3	US-09-677-064-10
4	384.8	25.1	1578	4	US-09-248-796A-3131
5	349.4	22.8	1231	3	US-09-397-787-34
6	349.4	22.8	1605	4	US-09-734-237B-74
7	334.6	21.8	1602	4	US-09-734-237B-72
8	161.4	10.5	3546	3	US-09-118-442-15
9	161.4	10.5	3546	3	US-09-677-064-15
10	159	10.4	3546	3	US-09-118-442-14
11	159	10.4	3546	3	US-09-677-064-14
12	92.6	6.0	77626	4	US-09-949-016-12608
13	63.8	4.2	294	4	US-09-313-294A-4684
14	44.2	2.9	1335	4	US-09-902-540-5256
15	44.2	2.9	3433	4	US-09-902-540-1257
16	38.8	2.5	7218	1	US-08-232-463-14
17	37	2.4	265038	4	US-09-949-016-15779
18	36.8	2.4	1141	4	US-09-806-708B-22
19	36.2	2.4	505	4	US-09-621-976-15639
20	36.2	2.4	568	4	US-09-370-767-1813
21	36.2	2.4	568	4	US-09-370-767-18095
22	35.6	2.3	1446	4	US-09-502-540-5772
23	35.6	2.3	72704	4	US-09-902-540-1273
24	35.4	2.3	274	3	US-09-118-442-21
25	35.4	2.3	3049	4	US-09-677-064-21
26	35.4	2.3	3049	4	US-09-300-958A-42
27	35.4	2.3	3049	4	US-09-949-016-991

28	35.4	2.3	3068	4	US-09-949-016-1297	Sequence 1297, Ap
29	35.2	2.3	57299	4	US-09-949-016-12141	Sequence 12141, A
30	35.2	2.3	153866	4	US-09-949-016-16919	Sequence 16919, A
31	35	2.3	315	4	US-09-270-767-1779	Sequence 1779, Ap
32	35	2.3	315	4	US-09-270-767-1779	Sequence 17061, A
33	35	2.3	600	4	US-09-669-751-171	Sequence 171, Ap
34	35	2.3	2025	4	US-09-107-532A-2547	Sequence 2547, Ap
35	34.8	2.3	1915	4	US-09-902-540-7256	Sequence 7256, Ap
36	34.8	2.3	6744	4	US-09-902-540-682	Sequence 682, App
37	34.6	2.3	1098	3	US-09-221-017B-928	Sequence 928, App
38	34.4	2.2	4085	3	US-09-165-240-5	Sequence 5, Appl1
39	34.4	2.2	4085	3	US-09-568-059-5	Sequence 5, Appl1
40	34.4	2.2	1664976	4	US-08-916-421B-1	Sequence 1, Appl1
41	34.4	2.2	1664976	4	US-09-692-570-1	Sequence 1, Appl1
42	34	2.2	1422	4	US-09-450-209-3	Sequence 3, Appl1
43	34	2.2	2941	4	US-09-949-016-2167	Sequence 2167, Ap
44	34	2.2	112112	4	US-09-949-016-15639	Sequence 15639, A
45	33.8	2.2	1362	4	US-09-328-352-116	Sequence 116, App

ALIGNMENTS

```
RESULT 1
US-09-727-628-1
; Sequence 1, Application US/09727628
; Patent No. 6791013
; GENERAL INFORMATION:
; APPLICANT: Armstrong, Katherine
; APPLICANT: Hey, Timothy D
; APPLICANT: Folkerts, Otto
; APPLICANT: Smith, Kelley A
; APPLICANT: Hopkins, Nicole L
; TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
; FILE REFERENCE: 50597
; CURRENT APPLICATION NUMBER: US/09/727, 628
; PRIOR APPLICATION NUMBER: 2000-12-01
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1959
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (137)..(1699)
US-09-727-628-1

Query Match      64.2%; Score 984.2; DB 4; Length 1959;
Best Local Similarity 77.6%; Pred. No. 0;
Matches 1190; Conservative 0; Mismatches 343; Indels 0; Gaps 0;

QY 1 AGTTCATGAGAAATTTTAAAGTAGAGACTCTTAATGTGAATACACCGAGCTAGATT 60
DB 137 AGTTCATGAGAAATTTTAAAGTAGAGACTCTTAATGTGAATACACCGAGCTAGATT 196
QY 61 CAGTCGAGTATACCACTGACCAACCGAATCTTGTTCAGGAACGAAATGACACTAT 120
DB 197 GAGTCGAGTATACCACTGACCAACCGAATCTTGTTCAGGAACGAAATGACACTAT 256
QY 121 CAGTCGAGTATACCACTGACCAACCGAATCTTGTTCAGGAACGAAATGACACTAT 180
DB 257 CGCTGGGTGTCGCGCCCAAGTCCGTCAGTCACTTCGAGCAGCAGCCGCGTCCC 316
QY 181 AATTCGGGGGTATGCTTGTGGGTGGAGTGAACCAACGCTTACCTTACCGGTGT 240
DB 317 AAGCTCGGGGTATGCTTGTGGGTGGAGTGAACCAACGCTTACCTTACCGGTGT 376
QY 241 GTTATTCCTAAGAGAGACATTTTCATGGGCTTAAAGGACAAAGATTTCAAGCAAT 300
DB 377 GTTATTCCTAAGAGAGACATTTTCATGGGCTTAAAGGACAAAGATTTCAAGCAAT 436
```

```
QY 301 TACTTGGCTCCCTCACCAGAGCTCAGCTATTGAGTTGATTCCTTCCAGGGAGAGAA 360
DB 437 TACTACGGCTCCCTCAACCCAGGCTTCCACATCAAGATGAGAGCTACAGGGAGAGAG 496
QY 361 ATCTATGCCCCATTCAGAGAGTGTCTTCCATGTTAATCTGACGACATGTGTGTTGGG 420
DB 497 ATATATGGCCGCTTCAAGAGCTCTCTACCATGTGTAACCCAGAGAGCATCTTGTGTTGGA 556
QY 421 GGATGGGATATCAGCAACATGMACTGCTGATGCTCATGAGCCAGGGCAAGAGTGTGTTGAC 480
DB 557 GGCTGGGACATCAGCAGATGAACCTGGCAGATGCGATACCAAGGCGCAAGGTGCTGAGC 616
QY 481 ATCGATTTGACAGAGAGATGAGGCTTACATGATTCATGATGCTTCCACTCCCGGAAATC 540
DB 617 ATTGACCTGACAGAGAGCTCAGGCCCCATCAGATGAGTCCATGCTGACCTTCCCGGTGTC 676
QY 541 TACGACCCGATTTTCAATGCTGCCAACCAAGAGAGCGTGCCAAACACGTGATTAAGAGGC 600
DB 677 TATGATCCGGACTTTCATGCGCGCTAACCGAGGCTCTCGTGCCAAACATGTTCATCAAGGCG 736
QY 601 ACAAGCAAGAGCAAGTTCAGCAATTCATCAAGACATCAGGCGCTTTAAGGAAGCCACC 660
DB 737 ACCAAGAAAGAACAGGTGAGAGATCATCAAGATATCAGGAGTTAAAGAGAAAGAAC 796
QY 661 AAAAGTGACAGAGGTGCTGCTGTCGACCAACAGAGAGGATAGCAATTTGGTT 720
DB 797 AAAAGTGACAGAGTATGCTGCTGTCGACCAACAGAGAGTACAGCAATGTATGCT 856
QY 721 GTAGGCTTTAATGACACATGAGAAATCTCTGCTGCTGTCGACAGAAATAGCGCTGAG 780
DB 857 GCTGGCTCAACGACACATGAGAAATCTGCTGCGCATCTGTGAGCAAGAAACAGCGCGAG 916
QY 781 ATTTCTCTTCCACCTTGATGCCATTTGCTGTCGATGAGAAAGTTCCTTCAATAT 840
DB 917 ATCTGCGCATCAACATATATGCAATTTGCTGTCAGAGAGGCTGCGTTCAATCAAT 976
QY 841 GGAAGCCCTCAGAAACCTTTGTACCAAGGCTGATGATCTTGCCATCCGAGAAACAT 900
DB 977 GGGAGCCCCCAAGAACCTTTTGTGCTGCGCTGATGATCTTGCTATCAAGAAACATCTGC 1036
QY 901 TTGATTTGAGAGATGACTTCAAGAGTGTGACAGCCAAATGAATCTGTGTGTTGAT 960
DB 1037 CTGATCGGTGAGAGACATTTCAAGAGTGGCAGAACCAAGATGAATCGTCTCGTTGAT 1096
QY 961 TTTCTTGTGGGGCTGTATCAAGCCAAATCTATGTATGATTAACAACATCTGGAAAC 1020
DB 1097 TTTCTTGTGTGTGCTGGAATAAAGCCACCTCGATGTAGCTAACCAACCTTGGGAAAC 1156
QY 1021 AATGATGATGATGATCTCGGCTCCAAACCTTCCGCTCCAAAGGAATCTCCAAAGAGC 1080
DB 1157 AACGAGCGATGATGATCTGCTGCTCCCTTCAAACTTATAGCTCAAGAGATCTTCCAAAGAGC 1216
QY 1081 AAGCTTTGTAGCATATGTGTCAACAGCAATGCCATCTCTATGAGCTGTGTGAATCC 1140
DB 1217 AAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1276
QY 1141 GACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
DB 1277 GATCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1336
QY 1201 GAGTACCTTTCAGAGATATTCATGAGGTGGAAGAAACACCATGTTTGTGCAACACATGT 1260
DB 1337 GAGTACACATCAGAGATCTTCAATGAGGCGGAGAGAGACATGTGTGTGTGTGTGTGTGT 1396
QY 1261 GAGGATTCCTTTTAACTGCTCTCTATATCTTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
DB 1397 GAGGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1456
QY 1321 ACTAAGATTCAGATTTAAAGCTGAAATGAGGAAATTTCCATCTCATTTCCACCATGTTGT 1380
DB 1457 ACCAGGATTCAGATTTAAACCTGAGGAGAGAGACCAATGTTCCATCTCTCTCTCTCTCTCTCT 1516
```

```
QY 1381 ACCATTTCAAGCTATCTGACCAAGGCTCTGTGTTCCACCGGGTACACCAAGTGTGAT 1440
DB 1517 ACCATCTTACCTACCTACCAAGGACCACTGTTCCACCGGACACCGGTGTGTAAC 1576
QY 1441 GCATTTGTAAGGACGCTGCAATGCTGGAATAAATAAGAGGCTGTGTGTGATTTGGCC 1500
DB 1577 GCTCTTGAAAGCAGAGGCGCATGCTGGAATAACATCAGAGGCTGTGTGTGCTGTGCC 1636
QY 1501 CCAGAGATTAACATGATCTGAGTACAAAGTGA 1533
DB 1637 CCAGAGAACACATGATCTGAGTACAAAGTGA 1669

RESULT 2
US-09-118-442-10
; Sequence 10, Application US/09118442B
; Patent No. 6197361
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytrate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 10
; TYPE: DNA
; LENGTH: 1931
; ORGANISM: Zea mays
US-09-118-442-10

Query Match 63.9%; Score 979.4; DB 3; Length 1931;
Best Local Similarity 77.4%; Pred. No. 5,8e-314;
Matches 1187; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

QY 1 ATGTTTCATCGAGATTTTAAAGTATGAGAGCTCTATATGTAAGTACACCGAGCTGAGATT 60
DB 99 ATGTTTCATCGAGAGCTTCCGCTCGAGAGCCCCCAGCTGCGGTACGCCCCGATGAGATC 158
QY 61 CAGTCGATGACCACTACGAAACACCGAACTTGTTCACGAGAAACAGAAATGCACTTAT 120
DB 159 GAGTCGAGTACCGGTATGACACGAGAGCTGTATACAGAGGAGCAAGAGCCGCTCTA 218
QY 121 CAGTGAATGTCMAACCCAAATCGTCAACTACAAATTTAAACCAACCAACCATGTTCCA 180
DB 219 CGCTGGGTGTCGCGCCCAAGTCGTCATCACTTCCGACCAAGACCGCGCTCCC 278
QY 181 AAATTTGGGGGTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
DB 279 AAGCTCGGGGTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 338
QY 241 GTTATTTGCTAACAGAGAGACATTCATGAGGCTACAAAGGACCAAGATTCACCAAGCCAT 300
DB 339 GTTATTTGCTAACAGAGAGAGAGATCTCATGTGGGACCAAGAGACATGTGTGACCAAG 398
QY 301 TACTTGGCTCCCTCACCAGAGCTCAGCTATTGAGTTGATTCCTTCCAGGGAGAGAA 360
DB 399 TACTACGGCTCCCTCAACCCAGGCTTCCACATCAAGATGAGAGCTACAGGGAGAGAG 458
QY 361 ATCTATGCCCCATTCAGAGAGTGTCTTCCATGTTAATCTGACGACATGTGTGTTGGG 420
DB 459 ATCTATGCGCCGTTCAAGAGCTCTTCCATATGTAACCCAGAGACATGTGTGTTGGA 518
```

QY 421 GATGGGATATCATGAACCTGGCTGATGCCATGGCCAGGGCAAGTGTGGAC 480
DB 519 GGCCTGGGACATTAAGCAATGAACCTGGCCGACTCCATGACGAGGGCAAGTGTGGAT 578
QY 481 ATCGATTTGAGAGAGCTTGAAGGCTTATCATGGAATCATGTGTTCCATCCCGGAATC 540
DB 579 ATTAGCCTGAGAGAGAGCTCAAGCCCTTATGAGAGTCCATGGTCCACTTCCCGGATC 638
QY 541 TACGACCCGAGATTCATGTGCTCCAAACCAAGAGAGCTCCAAACAGTGAATTAAGGC 600
DB 639 TATGATCCGAGCTTATCGCGCTTAAACAGGCTCTCCGCCCAAGTGTCAATCAAGGCG 698
QY 601 ACAAGCAAGAGAGAGTTCAGCAAAATCATCAAGACATCAAGCGCTTAAAGAGCCACC 660
DB 699 ACCAAGAAAGACAGGTGGAGAGATCATCAAGATATCAAGGAGTTTAAAGAGAAAGAAC 758
QY 661 AAGTGACACAGGTGTGTCTGTGGAAGTCCAAACAGAGAGATTAAGCAATTTGGTT 720
DB 759 AAGTGAGACAGATGATGTGTGTGGAAGTCCAAACAGTGAAGTATGCAATGTGTGC 818
QY 721 GTAGGCTTATATACACATGAGAAATCTTGGCTGTGTGGAAGAGAAATGAGCTGAG 780
DB 819 GCTGCTCTCAACGACAGATGAGAAATCTACTGCAATCTGTGGAAGAGAGAGGAG 878
QY 781 ATTTCTCTTCCACCTTGTATGACATGCTGTGTGATGGAAGAAATGTTCTTCTTATTAAT 840
DB 879 GTATCAACATCAACTATATGCAATGCTGTGTCTATGGAAGGGGTGCTTATCAAT 938
QY 841 GGAAGCCCTCAGAACTTTTGTACAGAGGCTGATTTGATTTGATCCATCCGAGGAACACT 900
DB 939 GGGAGCCCCAGAACACCTTTGTGCTGGCTGATTTGATTTGATTTTAAATAAACTGTC 998
QY 901 TTGATTTGAGAGATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTGAT 960
DB 999 TTGATTTGAGAGATGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTGTGAT 1058
QY 961 TTTCTTGGGGGGCTGTATCAAGCCCACTATAGTTAGTTCAACCATCTGGGAAAC 1020
DB 1059 TTCTCTTGTGTGTGTGAAATAAGCCCACTCAATGTGTGATCAACCACTTGGGAAAC 1118
QY 1021 AATGATGATGATGATCTCTCGGCTCCAAACCTTCCGCTCAAGGAAATCTTCAAGAG 1080
DB 1119 AAGAGTGGATGATGATCTCTCGGCTCCAAACCTTCAAGGATCTTCAAGAG 1178
QY 1081 AAGCTTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1179 AAGCTTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1238
QY 1141 GACATGTTGTTGTTATTAAGTATGTCCTTATAGTGGGATAGCAAGAGCATGAT 1200
DB 1239 GATCATGTGTGTGTCAAGATATGTCCTATAGTGGGATAGCAAGAGCATGAT 1298
QY 1201 GAGTACACTTCAAGATATTCATGAGTGGGAAAGACCAATTTTGTGCAACATGAT 1260
DB 1299 GAGTACACTTCAAGATATTCATGAGTGGGAAAGACCAATGTCGTGCAACATGAT 1358
QY 1261 GAGGATTCCTTTTAACTGCTCTATTAATCTTGAAGTGTGCTTGTGAGTGAAC 1320
DB 1359 GAGGATTCCTCTCCGCCCACTATCAATCTTGAAGTGTGCTTGTGAGTGAAC 1418
QY 1321 ACTAGAATCAAGTTTAAAGTGAAGTGAAGGAAATTCATCACTTCCACCAAGTTGCT 1380
DB 1419 ACCAGATCAAGTGAAGTGAAGTGAAGGAAATTCATCACTTCCACCGGAGTGC 1478
QY 1381 ACCATTTCAAGTATGACCAAGGCTCTCTGTGTGCAACCGGATGACCAAGTGTGAAT 1440
DB 1479 ACCATTTCAAGTATGACCAAGGCTCTCTGTGTGCAACCGGATGACCAAGTGTGAAT 1538
QY 1441 GCATTGTCAAGAGCGTGAATGCTGGAAGAAATGAAGGCTGTGTGTGATTTGAC 1500
DB 1539 GCTGTGGCAAGAGCGGCTGAATGCTGGAAGAAATGAAGGCTGTGTGTGATTTGAC 1598

QY 1501 CCAGAGATTAACATGATTCGATGACATGATG 1533
DB 1599 CCAGAGATTAACATGATTCGATGACATGATG 1631

RESULT 3
US-09-677-064-10
; Sequence 10, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phyrate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-10

Query Match 63.9%; Score 979.4; DB 3; Length 1931;
Best Local Similarity 77.4%; Pred. No. 5,8e-314;
Matches 1187; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

QY 1 ATGTTTCATGAGAAATTTAAGTGAAGAGTCTTATGTAAGTACCGAGACTGAGATT 60
DB 99 ATGTTTCATGAGAAATTTAAGTGAAGAGTCTTATGTAAGTACCGAGACTGAGATT 158
QY 61 CAGTCCGTGTAACATGAGAAACCAACCGGAACTTTTACAGAAACAGAAATGGCCATT 120
DB 159 GAGTCCGTGTAACATGAGAAACCAACCGGAACTTTTACAGAAACAGAAATGGCCATT 218
QY 121 CAGTCCGTGTAACATGAGAAACCAACCGGAACTTTTACAGAAACAGAAATGGCCATT 180
DB 219 CAGTCCGTGTAACATGAGAAACCAACCGGAACTTTTACAGAAACAGAAATGGCCATT 278
QY 181 AAATGGGGGTATGCTTGTGGGTTGGGGTGAACCAACGCTTACCTTACCGGTGCT 240
DB 279 AAGCTCGGGGTATGCTTGTGGGTTGGGGTGAACCAACGCTTACCTTACCGGTGCT 338
QY 241 GTTATTTGTAAGAGAGACATTTTATGAGGCTTACAAAGAGAAATTTCAAGCCATT 300
DB 339 GTTATTTGTAAGAGAGACATTTTATGAGGCTTACAAAGAGAAATTTCAAGCCATT 398
QY 301 TACTTTGGCTCCCTCACCACCAAGCTCAGCTATTCAGTGTGATCTTCCAGGAGAGAA 360
DB 399 TACTTTGGCTCCCTCACCACCAAGCTCAGCTATTCAGTGTGATCTTCCAGGAGAGAA 458
QY 361 ATCTATGCCCCATTAAGAGTGTGCTTCCATGTTAATCTTGAAGCAATGTGTTGG 420
DB 459 ATCTATGCCCCATTAAGAGTGTGCTTCCATGTTAATCTTGAAGCAATGTGTTGG 518
QY 421 GGAAGGATTAAGAGAACTGAACCTGCTATGAGCAATGAGCAAGGAGAAAGTGTGAC 480
DB 519 GGCCTGGCAATTAAGAACTGAACCTGCTATGAGCAATGAGCAAGGAGAAAGTGTGAC 578
QY 481 ATCGATTTGCAAGAGCAAGTGAAGCTTATGATGATGATGATGATGATGATGATGATGAT 540
DB 579 ATCGATTTGCAAGAGCAAGTGAAGCTTATGATGATGATGATGATGATGATGATGATGAT 638

```
QY 541 TACGACCCGATTTCATTGCTGSCCAACCAAGAGAGCGTGGCCAAACGATTAAGGCC 600
Db 639 TATGATCCGACTTCAATCGCGCTAACCAAGGCTCTCGGCCAAAGTCTATCAAGGCC 698
QY 601 ACMAAGCAGAGCAAGTTCCAGCAAAATCATCAAAAGCATCAGGCGTTTAAAGAACCCACC 660
Db 699 ACCAAGAAAGAACAGGTGAGAGAGATCATCAAGATATCAGGAGATTTAAGAGAAAGAAC 758
QY 661 AAAGTGGACAGAGTGTGTCTCTGTGAGATGCGCAACAGAGAGAGATATAGCAATTTGGT 720
Db 759 AAAGTGGACAGAGATGTGTGTGTGTGAGATGCAAAACATGAAAGGTATAGCATGTGTGC 818
QY 721 GTAGGCGCTTAATGACCAACATGAGAGAAATCTTGGCTGTGTGAGACAGAAATGAGCGTAG 780
Db 819 GCTGTGCTTCAACGACACAGATGAGAGATCTACTGTGCTGTGTGACAGAGAACAGCGGAG 878
QY 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGTATGAGAAATGTTCTTTCAATTAAT 840
Db 879 GTATCAACCATCAACCTATATGCAATTTGCTGTGTATGAGAAAGGCGTCCGTTCAATCAAT 938
QY 841 GGAAGCCCTCAAGACCTTTTGTACCAAGGCGTGTATGTATCTTGGCATCGCGAGAACACT 900
Db 939 GGGAGCCCCAGAACACCTTTGTGCTGTGCTGTATGTATCTTGTATTAATAAAACAACTGC 998
QY 901 TTGATTGTGAGATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTGTAT 960
Db 999 TTGATTGTGTGTGACACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTGTAT 1058
QY 961 TTTCTTTGGGGGCTGTGTATCAAGCCCAACATCTATAGTTAGTTACCAACATCTGGGAAAC 1020
Db 1059 TTCTTTGTGTGTGTGTGTATTAAGCCCACTCAATGTGTGTGTATCAACCACTTTGGGAAAC 1118
QY 1021 AATGATGTATTAATCTCTCGGCTCCAAACCTTCCGCTCCAAAGGAAATCTCCAAAGAC 1080
Db 1119 AACGATGTGATGAACCTGTCTGTGCTCCCAACATTAAGGTCCAAAGAGATCTCCAAAGAC 1178
QY 1081 AACGTTGTGACGATATGTGTCAACAGCAATGCCATCTCTATGAGCGCTGTGAACTCC 1140
Db 1179 AACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1238
QY 1141 GACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db 1239 GATCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1298
QY 1201 GAGTATCTTCAAGATATTTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db 1299 GAGTATCACTCAAGATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1358
QY 1261 GAGGATTTCCCTTTTACCTGCTCTATTTATCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Db 1359 GAGGATCTGCTCTCTGCGGCACTTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1418
QY 1321 ACTAAGATCCAGTTTAAAGCTGAAATGAGGAAATTCACCTCATTTCCACCCAGTTTGTCT 1380
Db 1419 ACCAGATCCAGCTGAAGCTGAGGAGAGAGAAATTCACCTCATTTCCACCCGAGTGTGCT 1478
QY 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
Db 1479 ACCATTTGTAGTTACTTCAACCAAGGCAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1538
QY 1441 GCATTTGCAAGAGCGTGTGATGTGTGAAACATATGAGAGGCTGTGTGTGTGTGTGTGTGT 1500
Db 1539 GCTCTGCGCCAGAGAGGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1598
QY 1501 CCAGAGATTAACATGATTTCTGAGTACAAAGTGA 1533
Db 1599 CCAGAGAACAAATGATCTTGTGAGTACAAAGTGA 1631
```

RESULT 4
US-09-248-796A-3131
; Sequence 3131, Application US/09248796A

```
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO: 3131
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Candida albicans
; US-09-248-796A-3131

Query Match 25.1%; Score 384.8; DB 4; Length 1578;
Best Local Similarity 56.0%; Pred. No. 1.4e-116;
Matches 812; Conservative 0; Mismatches 602; Indels 36; Gaps 3;
```

```
QY 110 ATGCACTATCATGATGTGTGTCAAAACCAATCCGTCAACTACCAATTTAAACCAACA 169
Db 113 ATGCTAATGTGTAAATTCATCTGTACCAACCACTGTAGTACTATGAAATTCAAAGTTGAT 172
QY 170 CCCATGTTCCAAATTTGGGGGATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 229
Db 173 TAAAGTCCCTTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 232
QY 230 TCACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 289
Db 223 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 292
QY 229 AACCAAGCAATTAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 343
Db 293 TTAACCAATTAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 352
QY 344 CCTTCCAGGAGAGGAAATCTATGCCCTTCAAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 403
Db 353 AAGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 412
QY 404 ACAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 463
Db 413 ATGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 472
QY 464 GGGCAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 523
Db 473 GAGCTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 532
QY 524 TTCCACTCCCGGAATCTACGACCCGGAATTTGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 583
Db 533 AACCTTTGGAATCAATCTTACCTGATTTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 592
QY 584 ACACGTGTATTAAG-----GGCAAAAGCAAGCAAGTTCT 619
Db 593 ACAATGTTTTTAACCAAGTCAATGTGTGAAGTTAAATCTATTAATTAATGAGCGGAGCTGTG 652
QY 620 AGCAATCTCAAAAGATCAAGGCTTTAAGAAAGCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGT 679
Db 653 AAAAAATCGAAAGATATCAAGATTTCAAGGCCAAAGCAAAATTAAGTAAAGTTATTA 712
QY 680 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 739
Db 713 TTTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 772
QY 740 TGAAGATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 799
Db 773 CTGCAACATTTGATCAAGATATTAAGATTCACAGAAATTTGTGTGTGTGTGTGTGTGTGTGT 832
QY 800 ATGCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 859
```


Db 833 TTGCGCTGCTTCTATCTTGAAAAAGTTCCATATATATAGTTCACCAACAAACACAT 892
Qy 860 TTGTACCAAGGCTGATGATCTTCCATGCGAGAAACACTTGTATGTGGAGATGACT 919
Db 893 TTGTTCCGAGTATATAGTGAAGTGAAGAAATACACATCATTCATTTGGTGGATGAT 952
Qy 920 TCAGAGGTGTGAGCAACAAATGAATCTGTGTGGTGAATTTCTTGTGGGGGCTGTA 979
Db 953 TCAGAGCAGGTCAACAAATTAATTAATCACTGTAGCTCAATTCCTTGTGATGCTGTA 1012
Qy 980 TCAGGCAACATCTATAGTATAGTATACACATCTGGGAAACATGATGATGATCTCT 1039
Db 1013 TCAGACCATCTTCTATGCTTCTTATATCACTTGGGTAACATGACGGTATCAATTTAT 1072
Qy 1040 CGGCTCAACAACTTCCGCTCCAGAGAAATCTCCAGAGCAACGTTTGTGAGATATGG 1099
Db 1073 CATCACAAACATTTAGATCTAGAGAAATTTCCAAACATCTGTGTGATGATATTA 1132
Qy 1100 TCAGACGACATGCGAT-----CCTCATGAGCGCTGTGAAACATCCGACCATGTTGTG 1153
Db 1133 TTGATATCAACGATATATGTATACACAGGAAATCTGTGACAAAGTTGATCACTGATCG 1192
Qy 1154 TTATTAAGATATGCTTATAGTATAGGAGATAGCAAGAGCATGATGATCACTTCAAG 1213
Db 1193 TCATTAATATCTTGCAGCGTGTGGATTTCTAAAGTTGCGATGATGATATCTATCAG 1252
Qy 1214 AGATATTCATGGGTGAAAGAACATGTTTTCACACACATGTGAGATTCCTTT 1273
Db 1253 AATTAATGTTGGGTGTCAACAAATTAATTAATCAATGTTTGTGAAGATTCATTCG 1312
Qy 1274 TAGCTGCTCTTATATCTTGAGAGTGTGCTCTGCTGAGCTGAGACATGAAATCCAGT 1333
Db 1313 TTGCTACACATGATGATTAATGATCTGTGTGGTGTGCTGATTAATTTGCACTAGCTCAG 1372
Qy 1334 TTAAAGCTGAAATGAGGAAATTCACATTCACCAAGTTGTACATTCCTGAGCT 1393
Db 1373 TCAGAGGCTCGGTAAATCTGATTAATGATGATGATGATCTGTGTGATTAATTTGCTT 1432
Qy 1394 ATCTGACCAAGGCTCTCTGTGTTCCACGGGTACACAGTGTGATGATGATCTCAAAG 1453
Db 1433 ACTGCTCAAGGCTCATTAAGCAAGCAAGATTCAAACCTATCAACGATTAACAAAC 1492
Qy 1454 AGGTCGCAATGCTGAAAAATTAATGAGGCTGTGTGATTTGGCCCAAGAAATACA 1513
Db 1493 AACGTCACAAATTAAGCAACCTGCTTCAAGTGTGTGATTAACATTAAGCAAGAT 1552
Qy 1514 TGATTTCTCGA 1523
Db 1553 TAAAGATTGA 1562

RESULT 5
US-09-397-787-34
; Sequence 34, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 1231
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-34

Query Match 22.8%; Score 349.4; DB 3; Length 1231;
Best Local Similarity 62.0%; Pred. No. 6.8e-105;
Matches 587; Conservative 0; Mismatches 356; Indels 4; Gaps 2;
Qy 584 ACAAGTGAATTAAGGCAACAAAGCAAGCAAGTTCAGAAATCATCAAGACATCAAG 643
Db 31 ACAACCTCATCCAGGCTCGCTGCGAGCACTGAGAGATATCCGACAGGACATCCAG 90
Qy 644 CGTTTAAGAACCCACCAAGTGAACAGGTGTGTCTGTGTGACCTGCAACACAGAA 703
Db 91 ACTTCGGCTCTAGCCGGGGCTGACAAAGTATGTGTGTGTGACGGGAAACAGAAC 150
Qy 704 GGTATAGCAATTTGGTGTGAGGCTTAATGACCAATGAGAAATCTTGGCTGTGTG 763
Db 151 GCTTCTGTGAGGTGATTCAGGCTTCACAGACACAGCGAAACCTGCTGCGACAT 210
Qy 764 ACAGAAATGAGGCTGAGATTTCTCTTCCACCTGTATGCCATTCGCTGTGTGATGAA 823
Db 211 AGCTGGGTCTG---GAGGTGTGCTCCAGCTCTGCGCGTGGCAGACATCTGAGAG 267
Qy 824 ATGTTCTTTTATTAATGAGAGCTTCAGAACATTTTGTATCCAGGCTGATTTG 883
Db 268 GCTGTGCTTCTCTCATATGAGGTCTCCGACAGAACCTGCTGCGAGCTTGTGAGCTG 327
Qy 884 CCATGCGCAGAAACCTTGAATTTGAGATGATGATCTTCAAGGTGTGACCAAAATGA 943
Db 328 CGTGGACACCGGTTTTTGTGGGGAGATGACTTCAAGTCAGGCCAGACCAAGTCA 387
Qy 944 AATCTGTGTGTGATTTCTTGTGGGGCTGTGTATCAAGCCAAATCTATATAGTTAGT 1003
Db 388 ACTCGGTGTGTGTGACCTTCATTTGCTCGGCTCAAGACATATCATGTGATGAT 447
Qy 1004 ACACCATGTGGGAAACATGATGATGAA-TCTCTGGCTTCACAAACCTTCCGCTC 1062
Db 448 ACACCATCTGGGCAACAGATGAGGAGAACCTTATGAGGCCCATTTGAGTTCCGCTC 507
Qy 1063 AAGGAAATCTCAAGAGCAAGTGTGACGATGATGATGATGATGATGATGATGAT 1122
Db 508 AAGGAGGTCTCCAGAGCAAGCTGTGGAGCAAGTGTGATGATGATGATGATGATGAT 567
Qy 1123 GAGCTGTGAAATCTCCGACCATGTTGTGTATTAATGATGATGATGATGATGAT 1182
Db 568 AGCCCGGAGAGAGCTGACACATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 627
Qy 1183 AGCAAGAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1242
Db 628 AGCAAGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
Qy 1243 GTTTGCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1302
Db 688 GTGTGCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 747
Qy 1303 CTTCCTGAGCTGAGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1362
Db 748 CTGCTGACCGAGCTGTGCAAGGCTGTGATGATGATGATGATGATGATGATGATGAT 807
Qy 1363 TCATTCACCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1422
Db 808 ACCTTCACCCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 867
Qy 1423 GGTACACAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1482
Db 868 GGCACCCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 927
Qy 1483 GCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1529
Db 928 GCTGCTGTGAGCTTCCGCGACAGAACATGATCTCTGGAACACAA 974

RESULT 6
US-09-734-237B-74
; Sequence 74, Application US/09734237B
; Patent No. 6818752

```
/ GENERAL INFORMATION:
/ APPLICANT: Rozzell, J. David
/ APPLICANT: Bui, Peter
/ APPLICANT: Hua, Ling
/ TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
/ FILE REFERENCE: B583:40608
/ CURRENT APPLICATION NUMBER: US/09/734,237B
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 09/494,921
/ PRIOR FILING DATE: 2000-01-31
/ NUMBER OF SEQ ID NOS: 79
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 74
/ LENGTH: 1605
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic gene derived from Saccharomyces cerevisiae myo-inositol
/ OTHER INFORMATION: -1-phosphate synthase, having numerous codons replaced with other
/ OTHER INFORMATION: s encoding the same amino acids to reduce free energy of folding,
/ OTHER INFORMATION: and a gly codon inserted after the initiating met codon
Us-09-734-237B-74
```

```
Query Match      22.8%; Score 349; DB 4; Length 1605;
Best Local Similarity 55.8%; Pred. No. 1.1e-104;
Matches 779; Conservative 0; Mismatches 565; Indels 51; Gaps 4;
```

```
QY 180 AAAATTGGGGGATGATCTGTGGGTTGGGTTGAAAACAAGGCTCTACCCGCGTGG 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 195 AAAGTGGGTATCATGCTATGCTGCGCTGGGTGTAACAAGGCTCTACTCTGTGCATC 254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 240 TGTATTGCTTAACAGAGAGACATTTTCATGGGCTACMAAGACAAGATTCAACAAGCCA 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 255 TGTTCGCGAAACAACAACAGTAGAATTCAGACTAAGAGAGGTTAAACAGCCGAA 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 300 TTACTTTGGCTCCCTCACCAGCCTCAGCTATTGAGTTGG---ATCCTTCAGAGAGA 356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 315 CTACTTTGGTTTATGACTCAGTGTCTACTCTGAAGCTGGGCATTGATCTGAAGGTAA 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 357 GGAATATCTATGCCCATTCAGAGTCTGCTCCAAATGTTAACTTCAGACATTTGTGT 416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 375 CGACGTTTACGCTCCGTTCACTCTCTGCTGCGATGATATCTCGAACGACTTGTGTGT 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 417 TGGGGATGGATATTCAGACATGAACTGTGCTGATGTCATGCGCAGGCAAGGTGT 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 435 TTCTGGTTGGATATCAACAGCGGATGTGTACAGAACATGACAGCGTTCTCAGGTCT 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 477 TGACATCGATTGTCAGAGACAGTGAAGCCTTAACATGAAATCCATGCTTCCACTCCCGG 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 495 GGAATATGATCTGCAACACGCTCTGAAGGCTAAGATGTCTGTGTTAAAGCCTGCGTC 554
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 537 AATCTACGACCGGATTTTATGCTGCGCAACCAAGAGGAGCGGCCAAC----- 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 555 CATCTACTAACCCGATTTTATGCACTTAACAGAGCGAACGCTTAACAACCTGATCAA 614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 586 -----AACGTATTAAAGGCAACAAGCAAGAGCAAGTTCAAGCAATCAT 629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 615 CCGTGAAGAAAAGGTAAGCTTACTACCGGTGTAAGTGAATCACTCAGCGATCGG 674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 630 CAAAGCATCAAGGCGTTTAAGAGCCCAAGATGACAAAGGTGTTGCTCTGTGAC 689
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 675 TCGTGTATTCAGAACTTCAAGAGAGAAAACGACATGGAACAAAGTATGTACTGTGGAC 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 690 TGGCAACAGAGAGGATATGCAATTTGTGTAGGCTTTAATGACACATGAGAGATGT 749
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 735 TGTAAACACTGAACGTTACGTAGAAATATCCCGGGTGAACGATACATGAGAAACCT 794
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 750 CTYGGCTGCTGAGACAGAAATGAGGCTGAGATTTCTCTCCACCTTGTATGCCATTGC 809
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 795 GGTGCAATATATCAAGACAGACAGAGAAATCGTCCGTCACCATCTTGGCTGCTGC 854
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 810 CTGTGTATGAGAAATGTTCTTTCTTATATGAAAGCCTCAGAACATTTTGTATCAGG 869
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
DB 855 ATCTATCTGGAAGCGTATCCGTACATCAAGGCTCTCCGAGAACACTTTGTAACGGG 914
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 870 GCTATTGATCTTCCATCGGAGAGAACACTTTGATTTGGTGAAGATGACTTCAAGATGG 929
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 915 TCTGGTACAGCTGGCTGAACAGAAAGTACTTATCGCTGTGACATCTGAATATCTGG 974
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 930 TGAACCAAAATGAAATCTGTGTTGTTGATTTTCTTGTGGGGCTGTATCAAGCCAAC 989
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 975 CCAGACTAACTGAAATCTGTACGTGACACAGTCTGTGTTACCGCTGTATCAACCGGT 1034
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 990 ATCTATGTTAGTTACAAACCATCTGGGAAAACAATGATGTGATATCTCGGCTCCACA 1049
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1035 TTCTATCGCTCTTATTAACCACTGGGTAAACAGACGCGCTACAACTGTCTGCTCGAA 1094
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1050 AACCTTCGCTCCAAAGAAATCTCCAGAGCAAGCTTGTGACATATGATCAACAGCA 1109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1095 AAGTTCGCTCTTAAGAAATCTTAATCTCTGTATGACGACATCATCGCTTCTTA 1154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1110 TGCCATCTCT-----CTATGAGCTGTGAAACATCCGACCATGTTGTTTATTAAGTA 1163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1155 CGACATCTGTACAAACGACAACTGGGTAAAGAAATGATCATCTGTATCGTTATCAATA 1214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1164 TGTGCTTACGTAGGAGATGACAGAGAGCCATGATGATGACATTCAGATATTCAT 1223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1215 CATGAACCGGTGTGATGATCTTAAGTTGCTATGACGAATTAATCTGAACTGATGCT 1274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1224 GGGTGAAGAAGAACCATTTGTTTGCACACATGTGAGGATTCCTTTTAGCTGCTCC 1283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1275 GGGGGGTACAAACGGTATCTATTCACAAAGTTTGTGAACATCTGTGCTGACACCC 1334
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1284 TATTATCTTGAATCTTGCTCTTCTTGTGCTGAGCTGACATGAAATCCAGTTAA----- 1337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1335 GCTGATCATGACACGCTGCTGTTATGACTGAATTTCTGATCCGCTGATCTTCAAGAAAGT 1394
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1338 -----AGCTGAAATGAGGGAATTCACATTCATTCACCCAGTTGCTACATCT 1388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1395 TGACCCGGTTAAAGAAATGCTGCGCAATTCGAAAACCTTACCCGGTTCTGACCTTCT 1454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1389 CAGCTATCTGACCAAGGCTCTCTGTGTTCCACCGGTAACACAGTGTGAATGCAATTGTC 1448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1455 GTCTTACTGCGTGAAGAGTCCGCTGACCTGTCCAGGCTTCCACCCGGTTAACGGTCAA 1514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1449 AAAGCAGGCTGATCTGGAACATATATGAGGCTTGTGATTTGGCCCCAGAGAA 1508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1515 CAAACAGGATACCGCTCTGGAACAACTTCTGCTGTGATGCGCCGCTCCAGAA 1574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1509 TAACTGATCTCGA 1523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1575 CGAAGTGGTTCGA 1589
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 7
Us-09-734-237B-72
/ Sequence 72, Application US/09734237B
/ Patent No. 6818752
/ GENERAL INFORMATION:
/ APPLICANT: Rozzell, J. David
/ APPLICANT: Bui, Peter
/ APPLICANT: Hua, Ling
/ TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
/ FILE REFERENCE: B583:40608
/ CURRENT APPLICATION NUMBER: US/09/734,237B
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 09/494,921
/ PRIOR FILING DATE: 2000-01-31
/ NUMBER OF SEQ ID NOS: 79
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 72
/ LENGTH: 1602
/ TYPE: DNA
/ ORGANISM: Saccharomyces cerevisiae
Us-09-734-237B-72
```

Query Match 21.8%; Score 334.6; DB 4; Length 1602;
 Best Local Similarity 54.2%; Pred. No. 6, 8e-100;
 Matches 847; Conservative 0; Mismatches 659; Indels 57; Gaps 6;

```

QY 17 TTAAGGTAGAGAGCTCTAATGTGAAGTACACCGAGCTGAGATTGAGTCGTCGCTGTAACACT 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 35 TTAAGGTAGATTCAGCAAGTGCACGTAACAGGACAAACGAGCTGCTCACCAGATACAGCT 94
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 77 ACAGAAACACCGAATCTTTGACAGAGAACGAAATGCGACCTATCGATGATTTGTAAC 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 95 ACAGAAATAGCTG---TAGTTACGAGACAGCTAGTGCCGCTTCGATGTAAACCCCACTG 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 137 CCAATCCCGTCACTACCAATTTAAACCAACCCATGTTCCAAATTTGGGGGTATG 196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 152 TTCAAGACTACGCTTTCAAACTTGACTTGAAAAAGCCG---AATAACTAGAAATTAATG 208
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 197 TTGTGGGTTGGGGTGAACCAACGCGCTTACCTCCACCGGTGTTGTTATTTGTAACAGAG 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 209 TCATTGGGTTAGGTGGCAACAATGGCTCCACTTAAGTGGCTGCGTATTTGGGCAATAGC 268
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 257 AGGACATTTTCATGCGCTACAAAGACAGATTCAACAGCAATTACTTTGGCTCCCTCA 316
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 269 ACAATGTGAGATTCAAACTTAAGAGAGCGGTTAAGCAACCAACTCTTCGCTCCATGA 328
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 317 CCCAAGCCTCAGCTATTGCGATTGATCCCTTC---AGGAGAGGAAATCTATGCCCAT 373
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 329 CTCATGTTCTACCTTGAAACCTGGGTATCGATCCGAGGGGAAATGACGTTATGCTCTT 388
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 374 TCAGAGCTGCTCCCAATGTTTATCTGACAGCAATTTGTTGGGGATGAGATATCA 433
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 389 TTAACCTCTGTTGCCATGTTTAGCCCAACGACTTTGTCGTCGTTGGTGGGACATCA 448
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 434 GCACATGAACCTGCTGATGCCATGCGCAAGGCAAAAGTGTGACATCGATTTCAG 493
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 449 ATTAACCCAGATCTATAGAACTATGCAAGATGCAAGATCTCGAATATGATCTCAAC 508
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 494 AGGAGTTGAGGCTTATCATGGAATCATGGTTCACATCCCGGAATCTACCGCGGAT 553
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 509 AACGTTGAAGGCGAAGATGTCCTTGGTAGCCCTCTTCATTTACTACCTCGAT 568
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 554 TCATTGCTGCCAACCAAGAGAGAGCTGCCAA-----CA 586
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 569 TCATTGCTGAGTATCAAGATGAGAGAGCAATACTGCATCAATTTGATGAAAAAGCA 628
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 587 AGCTGATTAAGGSCAACAAGCAAGCAAGTTGACAAATCATCAAGACATCAAGCGCT 646
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 629 AGCTAACCAAGAGGGGTAAGTGAGACCATCTGCACAGACCGCATATCCAGAAAT 688
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 647 TTAAGGAGCCACCAAGTGACAAAGTGAGTGTCTCTGAGATGCGCAACAGAGAGT 706
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 689 TCAGAAAGAAAAAGCCCTTGATTAAGTATCTTTGGACTGCAAAATACGAGAGT 748
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 707 ATAGCAATTTGGTTGAGGCTTATGACACCATGAGAAATCTCTGGGCTGCTGAGCA 766
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 749 AGCTAAGATATCTCTGCTGTTAATGACACCATGAAAACTCTTGCGAGTCTATTAAGA 808
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 767 GAAATGAGGCTGAGATTTCTCTTCCACTTGTATGCCATGCTGTGTGATGAAAAATG 826
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 809 ATGACCATGAAGAGATGCTCTCTTCCACGATCTTTCAGAGAGCATATCTTGAAGGTG 868
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 827 TTCCCTTCAATTAAGAGCCCTCAGAACACTTTTGTACAGGCGTGAATTTGATCTTGCCA 886
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 869 TCCCTATATTAATGTTGACCGCAAAATCTTTTGTCCCGCTGTGTTCACTGCGCTG 928
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 887 TCGCGAGAAACATTTGATTTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAT 946
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 929 AGCATAGGGTACATTCATTTGCGGAGAGCATCTCAAGTCGGGACAAACCAAGTTGAAGT 988
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 947 CTGTGTTGTTGATTTCTTGTGGGGCTGTGATCAAGCCCAATGTATAGTTAGTTACA 1006
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 989 CTGTTCTGGCCCGAGTTCTTAGTGAAGAGGTATTAACCGGCTCCTCATTTGATCTTATA 1048
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1007 ACCATCTGGGAAACATGATGATGATCTCTCGGCTCCACAAACCTTCGCTCCAAAG 1066
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 1049 ACCATTTAGCAATATAGACGGTTATTAATTTATCTGCTCCAAAAACAATTTAGTCTAAG 1108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1067 AAATCTCAAGAGCAACGTTGTTGACGATATGCTCAACGCAATGCCATCTCT----- 1120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1109 AGATTTCCAAAAGTTCTGTCTATAGATGACATCATCGCTCTATATGATATCTTGTAACATG 1168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1121 ATGAGCTGCTGTAACATCCCGACCATGTTGTTGTTATTAAGTATGTCCTTAAGTGGG 1180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1169 ATAAACTGGGTAAAAAGTTGACCACTGCAATTTGATCAAAATATTAAGAACCCGCTGGGG 1228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1181 ATAGCAAGAGCCCATGATGATGATCACTTCAGAGATATTCATGGGTGAAGAAACACCA 1240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1229 ACTCAAAAGTGGCAATGACGAGTATTAACAGAGATGATGTTAGTGGCCATAACCGGA 1288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1241 TTGTTTGAACAACAATGAGAGATTCCTTTTAACTGCTCTCTATTAATCTTGAACCTGG 1300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1289 TTTTCATTCACAATGTTTTCGAAGATTTCTTACTGGCTAGCCCTTGATCATGATCTTT 1348
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1301 TCCCTTCTGAGCTGAGCACTAGAAATCCAGTTTAAAGCT-----GAAA 1345
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1349 TAGTCAATGACTGATTTGTACAGAGTGTCTTATTAAGAGTGACCCAGTTAAAGAG 1408
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1346 ATGAGGAAATTTCCACTCATTCACCCAGTGTCTACATTCAGCTATTCGACCAAG 1405
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1409 ATGCTGGCAAAATTCAGAACTTTTATCCAGTTTAACTTTGAGTTACTGTTTAAAG 1468
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1406 CTCCTCTGTGCTCACCGGCTACACAGATGTAATGATTTGCAAGACGCTGCATGTC 1465
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1469 CTCATTTAACAAGACAGAGATTTCAACCCGGTGAATGGCTTAAACAAGAAAGACGGCT 1528
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1466 TGGAAAACTAATGAGGCTTGTGTTGATTTGGCCCCAGAGAAATAAGATTTCTGAGT 1525
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1529 TGAAGAAATTTTAAAGATTGTGATGATTCCTTCCAAAACGAACTAAGATTCGAG 1588
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1526 ACA 1528
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1589 ACA 1591
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 8
 US-09-118-442-15
 ; Sequence 15, Application US/09118442B
 ; Patent No. 6197561
 ; GENERAL INFORMATION:
 ; APPLICANT: Martino-Cat, Susan J.
 ; APPLICANT: Wang, Hongyu
 ; APPLICANT: Beach, Larry R.
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Bowen, Benjamin A.
 ; TITLE OF INVENTION: Genes Controlling Phytochemical Metabolism in
 ; FILE REFERENCE: 0706
 ; CURRENT APPLICATION NUMBER: US/09/118,442B
 ; CURRENT FILING DATE: 1998-07-17
 ; EARLIER APPLICATION NUMBER: 60/055,446
 ; EARLIER FILING DATE: 1997-08-11
 ; EARLIER APPLICATION NUMBER: 60/055,526
 ; EARLIER FILING DATE: 1997-08-08
 ; EARLIER APPLICATION NUMBER: 60/053,944
 ; EARLIER FILING DATE: 1997-07-28
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 15
 ; LENGTH: 3546
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; US-09-118-442-15

Query Match 10.5%; Score 161.4; DB 3; Length 3546;
 Best Local Similarity 64.5%; Pred. No. 4e-42;
 Matches 305; Conservative 0; Mismatches 84; Indels 84; Gaps 1;

QY	979	ATCAAGCAACATCTATGTAGTAGTACCAACATCTGGGAAACAATGATGATGATATTC	1038
Db	2716	ATGCAAGCCCACTCAATGTGAGCTACCAACCATTTGGGAAACAACGATGCGATGACCTG	2775
QY	1039	TCGGCTCCCAAACTTTCGCTCCAGGAAATCTTCAAGAGCAACGTTGTTGACGATATG	1098
Db	2776	TCTCCCTTCCAAACTTCAGGTCCAGAGAGATCTTCAAGAGCAACGTTGTTGATGACATG	2835
QY	1099	GTCAACAGCATGCGATCTCTATGAGCTGTGTAACATCCCGACCATGTTGTTGATTT	1158
Db	2836	GCTTCGACCAATGCCATCTCTATGAGCCCGGCGAGCATCCCGATCATGTGTTGTCATC	2895
QY	1159	A-----	1159
Db	2896	AAGGCTCTTAGCTGATCTTCACTCGTTAAAGTTGACATATGCAAGCGAGATTTACA	2955
QY	1160	-----AGTATGTGCTTACGTAGGGGATATGCAAGAGACC	1194
Db	2956	TTGAAACTGTGCACTTTTGTGTGCAAGTATGTGCGTACGTGGAGACAGCAAGAGGCT	3015
QY	1195	ATGATGATGATACATTGAGAGATATTAATGATGGTGAAGAAACAACATGTTTGCACAAC	1254
Db	3016	ATGACAGATACACTCGAGATCTTCAATGGGCGCAAGAACCAATGTGTCTGCACAAAC	3075
QY	1255	ACATGTGAGATTCCTTTTAGCTGCTCTATATATCTTGAATTGTGCTTCTTGTCTGAG	1314
Db	3076	ACCTGTGAGACTCGCTCTCGCGCACTATCATCTTGTATGTGTGCTCTTGGCTGAG	3135
QY	1315	CTGAGCATTAATTCAGCTTTAAAGCTGAANAATGAGGGAANAATTCATCTACTT	1367
Db	3136	CTCAGCAACAGATTCAGCTTAAGAGCTAAGGAGGTTAAGAGCCCTCCCAAGT	3188

```

RESULT 9
US-09-677-064-15
; Sequence 15, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
; TITLE OF INVENTION: Plants and User Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-15

```

	Query Match	Similarity	Score	DB 3	Length
Best Local	64.5%	10.5%	161.4	DB 3	3546
Matches	305	Conservative	0	Mismatches	84
				Indels	84
				Gaps	1

Qy	1099	GTCAACAGCAATGCCATCCCTCTATGAGCCGTGTGAAACATCCCAACCATGTTGTGTATT	1158
Db	2836	GTCTCGAGCAATGCAATCCCTCTATGAGCCCGGAGAGATCCCGATCATGTGTGTATC	2895
Qy	1159	A-----	1159
Db	2896	AAGCTCTGTTAGCTGATCTTCACTCTGTTAAAAAGTTGACATATGCAAGGACAGATTACA	2955
Qy	1160	-----AGTATGTGCTTTACGTAGGGATATGACAGAGACC	1194
Db	2956	TTGAAACTTGTCACTCTTTTGTGTGACATATGTGCGTACGTGGGAGACAGAAAGGGCT	3015
Qy	1195	ATGATGATGATCACTTATGAGATATTTATGTGGTGTGAAAGAACACACTTGTTTTGACAAAC	1254
Db	3016	ATGACACAGTATCACTCAGAGATCTTATGTGGCGGCAAGAACACACTGTGTGACAAAC	3075
Qy	1255	ACATGTAGGATTCCTTTAGCGTCCATTAATCTTGGACTGTGTCTTCTGTGTAG	1314
Db	3076	ACCTGTAGGACTGTGCTCTGTGCGGACATATCACTCTTGTATCTGTGGTCTTGTGCTGAG	3135
Qy	1315	CTGAGCACTGAATCCAGTTTAAAGCTGAAATATGAGGAAATTCACATCATTT	1367
Db	3136	CTGAGCACTGAATCCAGCTGAAGCTGAGGAGAGAGTTAAGAGCCGCCCAAGT	3188

```

RESULT 10
US-09-118-442-14
; Sequence 14, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118.442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-14

```

Query Match	10.4%;	Score 159;	DB 3;	Length 3546;
Best Local Similarity	64.3%;	Pred. No. 2.5e-41;		
Matches	304;	Conservative	0;	Mismatches 85; Indels 84; Gaps 1;
OY	979	ATCAAGCCACATCTTAAGTTAGTTATACACCATCTGGGAAACATGATGATTAACATCTC	1038	
Db	2716	ATGAGGCCCACTTAATCGTAGGTACACACACTTGGAAACAGATGGGATTAACCTG	2775	
OY	1039	TCGGCTCCACAAACCTTCGCTCCCAAGAAATCTCCAAGAGCAAGTGTGTGACGATATG	1098	
Db	2776	TCTCCCTTCAAAATTCAGGTCCAGGTCCAAGAGATCTCCAAGAGCAACGAGTGATGACATG	2835	
OY	1099	GTCACAGCAATGCATCTCTATAGACCTGGTAAACATCCCGACATGTTGTGTTATT	1158	
Db	2836	GTCCTGAGCAATGCCATCTCTATAGAGCCCGGACAGATCCCGATCATGTGTTGTATC	2895	
OY	1159	A-----	1159	

```

Db      2896 AAGGTCTGTAGCTGATCTTTGACCTCGTTAAAGTTGACATATGCAAGGACAGATTTACA 2955
Qy      1160 -----AGTATGTCCTTTAGCTGATGAGGATAGCAAGAGACC 1194
Db      2956 TTGAAACTTGTCACCTCTTTTGTGTGACATGTCCTGACCTGGGAGACAGCAAGAGGGCT 3015
Qy      1195 ATGAGATGATGACCTTCAGAGATATTCATGGGAGAAAGAACCACTGTTTGTGACAAAC 1254
Db      3016 ATGGACGATGACCTTCAGAGATCTTCATGGGAGGAGCAAGAACCACTGTCGACAAAC 3075
Qy      1255 ACATGTAGAGATTCCTCTTTAGCTGCTCTATATCTTGAGACTTGCTCTTCTGCTGAG 1314
Db      3076 ACCTGTAGAGACTCGCTCTCGCCGACCATCATCTTGATGTGGTGTCTTGCTGAG 3135
Qy      1315 CTGAGCACTGATGATCCAGTTTAAAGCTGAAATGAGGAGAAATTCACATCTATT 1367
Db      3136 CTCAGCACCAAGATCCAGCTGAAAGCTGAGGAGGGGTAAAGCCCCCAAGT 3188

```

RESULT 11
US-09-677-064-14
Sequence 14, Application US/09677064
Patent No. 6291224
GENERAL INFORMATION:
APPLICANT: Martino-Catt, Susan J.
APPLICANT: Wang, Hongyu
TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
TITLE OF INVENTION: Plants and Uses Thereof
FILE REFERENCE: 0706D
CURRENT APPLICATION NUMBER: US/09/677,064
CURRENT FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/055,446
PRIOR FILING DATE: 1997-08-11
PRIOR APPLICATION NUMBER: 60/055,526
PRIOR FILING DATE: 1997-08-08
PRIOR APPLICATION NUMBER: 60/053,944
PRIOR FILING DATE: 1997-07-28
PRIOR APPLICATION NUMBER: 09/118,442
PRIOR FILING DATE: 1998-07-17
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 3546
TYPE: DNA
ORGANISM: Zea mays
US-09-677-064-14

Query Match 10.4%; Score 159; DB 3; Length 3546;
Best Local Similarity 64.3%; Pred. No. 2.5e-41;
Matches 304; Conservative 0; Mismatches 85; Indels 84; Gaps 1;

```

Qy      979 ATCAAGCCAACATCTTATGTTACCAACATCTGGGAAACAATGATGATGATCTC 1038
Db      2716 ATGACGCCACCTCATCTGAGCTACCAACCACTTGGAAGAAACAATGATGATGATCTC 2775
Qy      1039 TCGGCTCCACAACTTCCGCTCCAGAAATCTCCAGAGCAACGTTTGTAGCATATG 1098
Db      2776 TCTGCCCTTCAAAACATTCAGGTCCAGAGAGATCTCCAAAGCAACGTTGTAGCATATG 2835
Qy      1099 GTCAACAGCAATGCTCTCTATGAGCTGTGGAACATCCGACCATGTTGTTTATT 1158
Db      2836 GTCTGAGCAATGCTCTCTATGAGCTGTGGAACATCCGATCATGTCTGTGATC 2895
Qy      1159 A----- 1159
Db      2896 AAGGTCTGTAGCTGATCTTTGACCTCGTTAAAGTTGACATATGCAAGGACAGATTTACA 2955
Qy      1160 -----AGTATGTCCTTTAGCTGATGAGGATAGCAAGAGACC 1194
Db      2956 TTGAAACTTGTCACCTCTTTTGTGTGACATGTCCTGACCTGGGAGACAGCAAGAGGGCT 3015
Qy      1195 ATGAGATGATGACCTTCAGAGATATTCATGGGAGAAAGAACCACTGTTTGTGACAAAC 1254

```

```

Db      3016 ATGACGAGTACCTTCAGAGATCTTCATGGGAGGAGCAAGAACCACTGCTGACAAAC 3075
Qy      1255 ACATGTAGAGATTCCTTTTAGCTGCTCTATATCTTGAGACTTGCTCTTGTGCTGAG 1314
Db      3076 ACCTGTAGAGACTCGCTCTCGCCGACCATCATCTTGATGTGGTGTCTTGCTGAG 3135
Qy      1315 CTGAGCACTGATGATCCAGTTTAAAGCTGAAATGAGGAGAAATTCACATCTATT 1367
Db      3136 CTCAGCACCAAGATCCAGCTGAAAGCTGAGGAGGGGTAAAGCCCCCAAGT 3188

```

RESULT 12
US-09-949-016-12608
Sequence 12608, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12608
LENGTH: 77626
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(77626)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12608

Query Match 6.0%; Score 92.6; DB 4; Length 77626;
Best Local Similarity 73.0%; Pred. No. 2.5e-18;
Matches 119; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```

Qy      983 ACCCAACATCTTATGTTACCAACATCTGGGAAACAATGATGATGATCTCTCG 1042
Db      34941 AGACCATGTCATCTGAGTTACAAACCTGGGCAACAAGATGGGAGAACTATCG 35000
Qy      1043 CTCACAAACTTCCGCTCCAGAAATCTCCAGAGCAACGTTGATGATGATGATCA 1102
Db      35001 CGCCATTCAGATTCGCTCTTAAAGAGGTGTCCAAAGAGCAACGTTGTGACGATGTGC 35060
Qy      1103 ACAGCAATGCTCTCTATGAGCTGTGGAACATCCGAGCA 1145
Db      35061 AAGGAAACCAAGTGTCTATACGCCCGGAGAGAGCTGACCA 35103

```

RESULT 13
US-09-313-294A-4684
Sequence 4684, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Lalgudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 4684

LENGTH: 294
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. 6476212 700348858H1
NAME/KEY: unsure
LOCATION: 185, 272, 275, 281, 283, 288
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4684

Query Match 4.2%; Score 63.8; DB 4; Length 294;
Best Local Similarity 65.7%; Pred. No. 1.8e-10;
Matches 92; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1 AGTTCATGAGAAATTTTAAAGTAGAGTCTTAATGTAGATACACGAGCTGAGATT 60
DB 125 ATGTTCAATGAGAGCTTCGCGTCGAGAGCCCGACGTCGGTACGCGCCGAGAGATC 184
QY 61 CAGTCGCTGACAACTACGAAACCAACCGAATTGTTACAGAGAAAGAAATGGCACTTAT 120
DB 185 NCTCGAGATACCGGATACGACACGAGAGCTGTGTCAAGAGCCAAAGACGCGCTCC 244
QY 121 CAGTGGATTGTCAACCCAA 140
DB 245 CGCTGGTGTCTCGGCCCAA 264

RESULT 14

US-09-902-540-5256
Sequence 5256, Application US/09902540
Patent No. 6833447

GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 5256
LENGTH: 1335
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-5256

Query Match 2.9%; Score 44.2; DB 4; Length 1335;
Best Local Similarity 44.3%; Pred. No. 0.0017;
Matches 322; Conservative 0; Mismatches 393; Indels 12; Gaps 3;

QY 374 TCAGAAGTCTGCTTCCAAATGTTAATCTGACGACATTTGTTGGGGATGGGATATCA 433
DB 194 TCACAGAGCTGGTGGCTTGGCGAGCTGAGAGACGTTGCTCGCGCTGGGACATCA 253
QY 434 GCAACATGAACCTGCGTGTATGCCATGCGCAAGGCAAGGTGTTGACATTCATTGCA 493
DB 254 TCCGTGAGAGCGCATACGAAGTGGCGTCCGCTCGGCGTGTCTCAAGCAAGCACTTG 313
QY 494 AGCAGTTGAGGCTTACATGAGATTCATGTTCCATCCCGGAATCTACGACCGCGATT 553
DB 314 AGAGGTGAAGCCGTTCTTCCAGAGATCAAGCCGAGAGAGGCGCTGCAAGACCTGAT 373
QY 554 TCATTGCTGCAACCAAGAGAGCGTGCACCAACCTGTATTAAAGGCAAAAGCAAGAC 613
DB 374 TGTGTGCGCGCATCGAGCGCAAC-----ACATCAAGGCAACCAAGAGCAACCGGAGA 427
QY 614 AAGTTCAAGAAATCATCAAGAGACATCAAGGCGTTTAAAGAGACCAACCAAGTGACAG 673
DB 428 GCATCGAAGCGCTGCGCCAGAGACATCCGGACTTCAAGAGAGCTCAACGCGAGCGCG 487

QY 674 TGTGTCCTGTGAGCTGCAACACAGAGAGTATAGCAATTTGTTAGGCTTAAATG 733
DB 488 CCGTATGTGTGTGTCAGACAGCGTGAAGACCTTCGTCGCGCTGCGAGTCTTCA--- 544
QY 734 ACACCATGAGAAATTTCTTGGCTGTGTGACAGAAATGAGCTGTAGATTTCTCTTCA 793
DB 545 AGACCCCTGCGCGCTTCCAGAAAGGCGCTGACAGAAAGCGCCGACATCAACCCACACG 604
QY 794 CTTGTATGCAATTCGCTGTGATGAAATGTTCTTTCAATTAATGAAAGCCCTCAGA 853
DB 605 CGCTGTACATTCACCGGCCATCAAGAGGCGGTGCTTTCGCAAGCCACGCCAACG 664
QY 854 ACACTTTTG---TACCAAGGCTGATTTGATCTTGCCATGCGAAGAACATTTGATG 910
DB 665 CCAAGGTGAGACACGCGCGCGCTCCAGAGATGCGCAACAGAGTCSGTGCGCGCGCG 724
QY 911 GAGATGACTTCAAGATGTGTACAGCAAAATGAAATCTGTGTGTTATTTCTTGTG 970
DB 725 GCCGAGACCTCAAGAGCGCGCAGAGATGATGAACGCTATTCGCGCGCTCAAGG 784
QY 971 GGGCTGTATCAAGCCAACTATATAGTTAATTACAAACATCTGGGAAACAATATG 1030
DB 785 CCGCATGCTGGGTTGTGATGCTGTGTTCTCCACCAATCTGAGCAACCGCAGCG 844
QY 1031 TGAATCTCTCGGCTTCCAAACCTTCCGCTCCAGGAAATCTCAAGAGCAACGTTG 1090
DB 845 AGGTGCTGAGAGACCCCGCGCTTCAAGGCGCAAGAGTCAACAAATGCAAGCGTGTG 904
QY 1091 ACGATAT 1097
DB 905 ACACCAT 911

RESULT 15

US-09-902-540-1257/c
Sequence 1257, Application US/09902540
Patent No. 6833447

GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 1257
LENGTH: 34316
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-1257

Query Match 2.9%; Score 44.2; DB 4; Length 34316;
Best Local Similarity 44.3%; Pred. No. 0.017;
Matches 322; Conservative 0; Mismatches 393; Indels 12; Gaps 3;

QY 374 TCAGAAGTCTGCTTCCAAATGTTAATCTGACGACATTTGTTGGGGATGGGATATCA 433
DB 19705 TCACAGAGCTGGTGGCTTGGCGAGCTGAGAGAGTGTGCTTGGCGCTGGGACATCA 19646
QY 434 GCAACATGAACCTGCGTGTATGCCATGCGCAAGGCAAGGTGTTGACATTCATTGCA 493
DB 19645 TCCGTGAGAGCGCATACGAAGTGGCGGTGCTCGGCGTGTCTCAAGCAAGCACTTG 19586
QY 494 AGCAGTTGAGGCTTACATGAGATTCATGTTCCATCCCGGAATCTACGACCGCGATT 553
DB 19585 AGAGGTGAAGCCGTTCTTCCAGAGATCAAGCCGAGAGAGGCGCTGCAAGACCTGAT 19526
QY 554 TCATTGCTGCAACCAAGAGAGCGTGCACCAACGATATTAAAGGCAAAAGCAAGAGC 613

```
Db 19525 TCGTGGCCGCAATCGAGGCCAAC-----ACATCAAGGCCACCAAGACCGCACCGGAGA 19472
QY 614 AAGTTAGCAAAATCATCAAGACATCAAGCGCTTTAAGAAAGCCACCAAGTAGACAAAG 673
Db 19471 GCATCCAAAGCGCTCGCCAGGACATCCGGCATTTCAAGAAAGAGCTCAACGGCAGCGCG 19412
QY 674 TGGTTGCTCGTGGAGCTGSCAACAGAGAGGATAGCAATTTGGTTGAGGCTTAATG 733
Db 19411 CCGTGAATGGTGGTGTGAGAGAGCGTGAAGACCTTCGTCGCTGCCGAGTCTTCA--- 19355
QY 734 AACACATGGAAGATCTTGTGCTGTGTGACAGAAATGAGGCTGAGATTTCTCTTCCA 793
Db 19354 AGACGCTGGCCGCTTCGAGAAAGCGCTGAGACGAAACAGCCCGACATCAACCCACCG 19295
QY 794 CCTTGTATGCAATGCTGCTGTGTGATGAAATGTTCTTTCAATTAATGAAAGCCTCAGA 853
Db 19294 CGCTGTACACCTACGCGGCATCAAGAGAGGCGTGCCTTCGGAAAGCCACGCCACAG 19235
QY 854 AACCTTTG---TACGAGGGCTGATGATCTTGCCATCGGAGGAACATTGATTGTG 910
Db 19234 CCAAGCTGACACAGCGCGGCTCCAGAGAGATGGCCAAGCAGAGTGGCGGTGCGCG 19175
QY 911 GAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGATTTTCTGTGG 970
Db 19174 GCCGCACTTCAAGAGCGGCGACAGATGATGAAGAGGTCAATCGCGCCCGCTCAAG 19115
QY 971 GGGCTGGTATCAAGCCAAACATCTATAGTTATACAACCATCTGGGAAACAAATGATGTA 1030
Db 19114 CCCGCAATGCTGGGTCTGATGAGTGTGTTCTCCACCAACATCTGGGCAACCGGAGCGCG 19055
QY 1031 TGAATCTTCGGGTCCCAAACTTCCGCTCCAGGAATCTTCAAGAGCAAGCTTGTG 1090
Db 19054 AGGTGCTGACGACCCCGCGCTTCAAGGCCAAGGAAGTCAACCAAGTCGAGGCTGTG 18995
QY 1091 ACGATAT 1097
Db 18994 ACACCAT 18988
```

Search completed: June 8, 2005, 03:03:36
Job time : 274.095 secs

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 22:54:36 ; Search time 5222.81 Seconds
(without alignments)
11172.621 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533
Sequence: 1 atgttcacgcaggaattttaa.....tgattctgcagracaaagta 1533

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_esc4:*
5: gb_esc5:*
6: gb_esc6:*
7: gb_esc7:*
8: gb_esc8:*
9: gb_esc9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	980.2	63.9	1740	9	CL961092 OeIFCC005
2	810.6	52.9	1628	3	CK827819 Arabidops
3	769.4	50.2	813	7	CK768601 Gm-r1030-
4	766.2	50.0	2360	1	AJ583520 AJ583520
5	698.8	45.6	988	7	CK277950 EST724028
6	651.4	42.5	914	7	CK275352 EST721430
7	648.4	42.3	936	7	CK279064 EST725142
8	636.2	41.5	865	7	CO085839 GR_Ba03B
9	633	41.3	906	7	CK320178 L2P08a02
10	628.2	41.0	898	7	CO113870 GR_Bb014
11	625.8	40.8	670	4	BG044525 BA029607
12	624.4	40.7	860	7	CO072025 GR_Ba30P
13	623.6	40.7	842	7	CF211061
14	623.4	40.7	867	6	CB973407 CAB30002
15	620.4	40.5	891	7	CF212628 CGF100065
16	616	40.2	836	7	CF518404 CAP0007_I
17	610.2	39.8	618	2	AM279066 BG07402.Y
18	606	39.5	734	6	CA801137 SA010109
19	597.2	39.0	746	6	CA918410 EST642557
20	593	38.7	594	4	BM523521 sam85a05
21	592	38.6	782	2	AW348136 AW348136
22	591.2	38.5	860	7	CO082095 GR_Ba46I
23	590.2	38.5	762	2	BE660321 1083 Gmax
24	588.6	38.4	707	2	BF005128 EST433626

c	25	587	38.3	754	4	BI968101	BI968101 GM830004A
c	26	586.8	38.3	834	7	CK933094	CK933094 CGF100434
c	27	585.4	38.2	978	4	BG388516	BG388516 GC01_01h1
c	28	583.2	38.0	831	7	CO132349	CO132349 GR_EB45C
c	29	582.2	37.9	672	2	AM508269	AM508269 B151h10.Y
c	30	581.2	37.9	883	6	CD438045	CD438045 EL01N0508
c	31	581	37.9	960	7	CK274458	CK274458 EST720536
c	32	580.4	37.9	622	2	AM397496	AM397496 BG79612.Y
c	33	575.4	37.5	830	7	CO107650	CO107650 GR_EB003
c	34	574.8	37.5	841	7	CO075472	CO075472 GR_Ea36C
c	35	573.6	37.4	816	6	CB975215	CB975215 CAB30006
c	36	573.2	37.4	663	6	CA900542	CA900542 PSC17211
c	37	573.2	37.4	891	7	CV290142	CV290142 acf01-10m
c	38	570.8	37.2	769	6	CA900545	CA900545 PSC19118
c	39	568.8	37.1	680	4	BG452447	BG452447 NF081D04L
c	40	563.6	36.8	670	2	BE191260	BE191260 Bn94D09.Y
c	41	560.6	36.6	685	2	BF634972	BF634972 NF077A08D
c	42	560.4	36.6	648	1	AT794750	AT794750 BB68E07.Y
c	43	560.4	36.6	713	4	BI111216	BI111216 EST511296
c	44	557.6	36.4	585	4	BG239011	BG239011 Bdb57d09
c	45	557	36.3	509	6	CB345158	CB345158 CA48BN000

ALIGNMENTS

RESULT 1
LOCUS CL961092 1740 bp DNA linear GSS 21-SEP-2004
DEFINITION OeIFCC005463 Oryza sativa Expressed Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
ACCESSION CL961092
VERSION CL961092.1 GI:52376905
KEYWORDS GSS.

SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Bukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae;

REFERENCE 1 (bases 1 to 1740)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,

Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. M. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)

JOURNAL
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES
source Location/Qualifiers
1..1740
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 63.9%; Score 980.2; DB 9; Length 1740;
Best Local Similarity 77.6%; Pred. No. 2,7e+285;
Matches 1186; Conservative 0; Mismatches 343; Indels 0; Gaps 0;

QY	1	ATGTTTCATCGAGATTTTAAGGTAGAGTCTTAATGAGTACACCGAGCTAGATT	60
DB	1	ATGTTTCATCGAGAGCTTCGCGTGGAGAGCCGCGTACGCGCGCGGAGATC	60
QY	61	CAGTCCGTGTACACTAGAAACACCGAAGCTTGTTCACGAGAAAGAGATGCACTAT	120

Db 61 GAGTGGAGTACAGACAGACAGAGAGCTGTGTGACAGAGACAGACAGAGCGCCCTCC 120
 Qy 121 CAGTGGATTTGTCAAAACCCAAATCCGTCATCAATTTAAACCAACCAACCATATGTCA 180
 Db 121 CGCTGGGTGTCTCGCCCAAGTCCGTCCGTCAACTTCGACCAACCAACCGTCC 180
 Qy 181 AAATGGGGGTATGCTTGTGGGTGTGGAGAAACAAGCTCTTACCTCTACCGGTGT 240
 Db 181 AAGCTCGGGGTATGTCTGTGGGTGTGGAGCAACAAGCTCTCAACGCTGACGGCTGG 240
 Qy 241 GTTATTTGCTAACAGAGAGACATTTTCATGTGGCTAACAGAGACAAATTCACAAAGCAAT 300
 Db 241 GTTCATGCGCAAGAGAGGGAATCTCATGTGGAGACCAAGAGACAAAGTGCAGCAACCAAC 300
 Qy 301 TACTATGTGCTCACTACCCAGAGCGTCCATCAGGGGTGGAAGCTACAAACGGGAGAG 360
 Db 301 ATCTATGCCCCCATTCAGAGTCTGTCCATGTGTTAATCTGACGACATTTGTGTTGG 420
 Qy 361 ATCTHAGCGCCCTTCAAGAGCTCTGTCCCATGTGTGAACCTTGATGACCTTGTGTGG 420
 Db 421 GGATGGGATATCAGCAACATGAACTGTGCTGATGCCATGCGCAGAGGCAAAAGTGTGAC 480
 Qy 481 ATCGATTTGCAAGAGAGTGGAGCGCTTTCATGGAATCCATGGTTCCATCCCGGAAATC 540
 Db 481 ATTTGATCTGCAAGAGAGCTCAGACCTTTCATGAGTCCATGTGCTCTCCCGGCAATC 540
 Qy 541 TAGCAGCCCGATTTTCATGCTCTGCCAAACAAAGAGAGCTGCCAAACACGTGATTAAGAG 600
 Db 541 TATGAGCCCGACTTCATGCGCCGCAACAGAGATCCCGGCAACAAATGTCTCAAGAGGC 600
 Qy 601 ACAAAGCAAGACCAAGTTCAGCAATCATCAAAAGACATCAAGGCGTTTAAAGAGCCACC 660
 Db 601 ACCAAGAGAGACAGATGAGAGACATCATCAAGACATCAGGAGAGTTCAAGAGAAAGAGC 660
 Qy 661 AAGAGGAGAGAGTGTGTCTCTGTGAGATGCGCAACAGAGAGATGCAATTTGTT 720
 Db 661 AAGAGGAGAGAGTGTGTCTCTGTGAGATGCGCAACAGAGAGATGCAATTTGTT 720
 Qy 721 GTAGGCTTAATGACACCATGAGAGATCTTGTGCTGTGTGAGACAGAAATGAGCTGAG 780
 Db 721 GTTGGGCTAATGACACATGAGAGACCTCTGTGGCTGTGTGAGACAGAAAGAGGAG 780
 Qy 781 ATTTCTCTTCACTTGTATGACATGCTGTGTGTGAGAGAAATGTTCTTTCAATAT 840
 Db 781 ATATCAACCATGACACTGTATGCAATGCTGTGTGTGAGAGAGGTATATACCGTTCAATTAAC 840
 Qy 841 GGAAGCCCTCAGAAACCTTTTGTACAGAGGCTGTATGATCTTGCCATGCGAGAGCAACT 900
 Db 841 GGAAGCTCTCAGAAACCTTTTGTGCTGTGAGCTGTATGCTTATTAAGAACAACTGCG 900
 Qy 901 TTGATTTGTGAGAGTACTTCAAGAGTGTGTGAGACCAAAATGTAATCTGTGTTGTTGAT 960
 Db 901 CTGATTTGTGTGTATGATTTCAAGAGTGTGAGACAAATGTAATGTAATCTGTGTTGAT 960
 Qy 961 TTTCTTTGTGGGGCTGTGTATCAAGCCACATCTAATGTTAGTTACACCATCTGGAAAC 1020
 Db 961 TTCTCTGTTGTGTGTGAGATTAAGCCACCTCAATGTGAGCTACAAACCACTTTGGGAAAT 1020
 Qy 1021 AATGATGATGATGATATCTCGGCTCCAGAAACCTTCCGCTCCAGAGAAATCTCCAAAGAGC 1080
 Db 1021 AATGATGATGATGATATCTCGGCTCCAGAAACCTTCCGCTCCAGAGAAATCTCCAAAGAGC 1080
 Qy 1081 AAGCTTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 Db 1081 AATGAT 1140
 Qy 1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGAGATGACAAAGAGAGCCATGAT 1200

Db 1141 GATCATGTTGTGTGATCAAGTATGTGCCGTATGTGGAGACAGACAAAGGCAATGAC 1200
 Qy 1201 GAGTACACTTCAGAGATATTCATGAGTGTGAAGAAACACCATTTGTTTTCACAAACATATG 1260
 Db 1201 GAGTACACTTCAGAGATATTCATGAGTGTGAAGAAACACCATTTGTTTTCACAAACATATG 1260
 Qy 1261 GAGATTCCTTTTATGCTGTCTTATTTATCTTGAATTTGTGCTTCTTGTGAGCTGAGC 1320
 Db 1261 GAGACTCACTTCTTCCGCGCCGATCATTTCTGATCTGTGTCTTCTTCCGAGCTGAGC 1320
 Qy 1321 ACTAGAAATCCAGTTTAAAGCTGAAGAAATGAGGAAATTCCTCATATCCACCAAGTGTCT 1380
 Db 1321 ACCAGATTTCACTGAAGCCGAGGGGAGAGAAATTCATTTCTTCATCTCATGTGTCT 1380
 Qy 1381 ACCATTTCACTGATATCTGACCAAGGCTCTGTGTTCCAGCGGTACACCATGTGTGAT 1440
 Db 1381 ACCATCTGAGTACTTCAACCAAGGACCTTGTCTCTCTGAGACACCATGTGTGAT 1440
 Qy 1441 GCATTTGCAAAAGCAGCGTGCATATGCTGAAAAATATATGAGGCTTGTGATTTGCTC 1500
 Db 1441 GCCCTGGCAAAAGCAGAGGCAATGCTTGAGAAACATCATGAGAGGCGCTGTGGCTGGCC 1500
 Qy 1501 CCAAGAAATACATGATTTCTGAGTACAA 1529
 Db 1501 CCGAGAAACAATGATCTTGAATACAA 1529

RESULT 2
 CDS04M4K
 LOCUS
 DEFINITION
 Arabidopsis thaliana full-length cDNA complete sequence from clone GSUTPGH24E06 of Hormone Treated Callus of strain col-0 of Arabidopsis thaliana (thale cress).
 BX827819
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1828)
 REFERENCE
 AUTHORS
 Castelli V., Aury J.M., Jallion O., Wincker P., Clepet C., Menard M., Cruaud C., Querlet F., Scarpelli C., Schachter V., Temple G., Caboche M., Weissenbach J. and Salanoubat M.
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with phrap.
 http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full1_length
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
 FEATURES
 source
 Location/Qualifiers
 1..1828
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"

software (<http://www.tigr.org/software/>).
 This clone was originally generated by the Public Soybean EST
 Project (<http://129.186.26.94/soybeanest.html>)/Shoemaker, R
 (reshnee@iasbate.edu).
 This clone is available through: Biogenetic Services, 801 32nd Ave,
 Brookings, SD 57006. For further information call 605-697-8500 or
 contact info@biogeneticservices.com
 Seq primer: T7-1 (5' AAT ACG ACT CAC TAT AG 3').
 Location/Qualifiers

FEATURES

source

1..813
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="Gm-r1030-621"
 /lab_host="DH10B"
 /clone_lib="Gm-r1030"
 /note="vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
 cDNA library was constructed from mRNA isolated from
 immature cotyledons of greenhouse grown plants
 (individual seed fresh weight of 100-300mg). The library
 was prepared using the life Technologies pSuperScript cDNA
 library construction kit. Complementary DNA was
 synthesized from mRNA using a poly(dT) sequence with a
 NotI restriction site. SalI linkers adapters were ligated
 to the blunt-ended cDNA fragments followed by NotI
 digestion. The cDNA fragments were directionally cloned
 into the NotI-SalI restriction site of the pSPORT1
 vector. The ligated cDNA fragments were transformed into
 E. coli Electromax DH10B host cells. This library was
 constructed by Dr. Lila Vodkin and Dr. Anu Khanna. Note
 that Gm-r1030 is a re-rack of Gm-cl007."

ORIGIN

Query Match 50.2%; Score 769.4; DB 7; Length 813;
 Best Local Similarity 97.4%; Pred. No. 1.9e-221;
 Matches 782; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

432 CAGCAACATGAACCTGGCTGATGCCACGAGGCAAGTGTTCATCATGATTGCA 491
 11 CGGCACATGAACCTGGCTGATGCCACGAGGCAAGTGTTCATCATGATTGCA 70
 492 GAAGCAGTTGAGGCTTACATGAAATCATGTTCCCTCCGGAATCTACGCCGA 551
 71 GAAGCAGTTGAGGCTTACATGAAATCATGTTCCCTCCGGAATCTACGCCGA 130
 552 TTTCATGTGTCACCAAGAGAGGCTGCCAACAACGATTTAAGGACCAAGCA 611
 131 TTTCATGTGTCACCAAGAGAGGCTGCCAACAACGATTTAAGGACCAAGCA 190
 612 GCAAGTTCAGAAATCATCAAGAATCAAGGCGTTTAAAGAACCAACAAAGTGACA 671
 191 GCAAGTTCAGAAATCATCAAGAATCAAGGCGTTTAAAGAACCAACAAAGTGACA 250
 672 GGTGGTGTCTGTGAGCTGCCAACAAGAGATATGCAATTTGGTTAGGCTTAA 731
 251 GGTGGTGTCTGTGAGCTGCCAACAAGAGATATGCAATTTGGTTAGGCTTAA 310
 732 TGAACCATGAGAAATCTTGGCTGCTTGACAGAAATGAGCGTGAATTTCCCTTC 791
 311 TGAACCATGAGAAATCTTGGCTGCTTGACAGAAATGAGCGTGAATTTCCCTTC 370
 792 CACCTTGATGCAATGCGCTGTGTGATGAGAAATGTTCTTTCATTTAAGAACCTCA 851
 371 CACCTTGATGCAATGCGCTGTGTGATGAGAAATGTTCTTTCATTTAAGAACCTCA 430
 852 GAACACTTTTGTACACGAGGCTGATTTGATCTTGCATCGCAGAGAACACTTTGATGG 911
 431 GAACACTTTTGTACACGAGGCTGATTTGATCTTGCATCGCAGAGAACACTTTGATGG 490
 912 AGATGACTTCAAGAGTGTCAAGCAAAATGAAATCTGTGTGATTTCTTGTGG 971
 491 AGATGACTTCAAGAGTGTCAAGCAAAATGAAATCTGTGTGATTTCTTGTGG 550

QY 972 GCGTGATATCAAGCAACATCTATAGTTAGTTACACCATCTGGGAAACAATGATGTTAT 1031
 DB 551 GCGTGATATCAAGCAACATCTATAGTTAGTTACACCATCTGGGAAACAATGATGTTAT 610
 QY 1032 GAATCTCTGGGCTCCACAAACCTCCGCTCAAGAAATCTCCAAAGCAAGTGTGTA 1091
 DB 611 GAATCTCTGGGCTCCACAAACCTCCGCTCAAGAAATCTCCAAAGCAAGTGTGTA 670
 QY 1092 CGATATGGTCAACAGCAATGCGCATCTATAGAGCTGTGTGAGATCCGACCATGTTGT 1151
 DB 671 TGAATATGATCAACAGCAATGCGCATCTATAGAGCTGTGTGAGATCCGACCATGTTGT 730
 QY 1152 TGTATTAATGATGTGCTTACGTAAGGAGATGACAGAGACCATGATGATCATCTTC 1211
 DB 731 TGTATTAATGATGTGCTTACGTAAGGAGATGACAGAGACCATGATGATCATCTTC 790
 QY 1212 AGAGTATTTCAATGGGTGGAAGA 1234
 DB 791 AGAGTATTTCAATGGGTGGAAGA 813

RESULT 4

AJ583520/c 2360 bp mRNA linear EST 08-JAN-2004
 LOCUS AJ583520 PpProt1 Physcomitrella patens cDNA clone pp001015066
 DEFINITION similar to putative myoinositol-1-phosphate synthase, mRNA
 sequence.

ACCESSION AJ583520
 VERSION AJ583520
 KEYWORDS GI:40781590
 SOURCE EST.

ORGANISM Physcomitrella patens
 Physcomitrella patens

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
 1 (bases 1 to 2360)
 Hohe,A., Egner,T., Lucht,J.M., Holterf,H., Reinhard,C., Schween,G.
 and Reek,R.

TITLE An improved and highly standardised transformation procedure allows
 efficient production of single and multiple targeted gene-knockouts
 in a moss, Physcomitrella patens

JOURNAL Curr. Genet. 44 (6), 339-347 (2004)

COMMENT Contact: Schween G
 Plant Biotechnology
 University of Freiburg
 Sonnenstrasse 5, D-70104 Freiburg, Germany.

FEATURES

source

1..2360
 /organism="Physcomitrella patens"
 /mol_type="mRNA"
 /db_xref="taxon:3218"
 /clone="Pp001015066"
 /tissue_type="Protonema"
 /clone_lib="PpProt1"

ORIGIN

Query Match 50.0%; Score 766.2; DB 1; Length 2360;
 Best Local Similarity 70.3%; Pred. No. 2.8e-220;
 Matches 1088; Conservative 0; Mismatches 443; Indels 16; Gaps 4;

1 ATGTTCAATGAGAAATTTAAGGTAGAGAGTCTTAATGTGAAGTACACGAGACTAGATT 60
 2150 ATGTTGTCGACAGCTTTCGTTGAGAGTCTCTCATGTGGAATACACGAGAGCATGATC 2091
 DB 61 CAGTCGGTGTCAACTAGAAACACGGAACCTGTTCAGAG-----AACGAAATGGC 114
 DB 2090 AAGTGAAGCTATGAGTACCAACACGAGTTCCTGCACGACGACGCAAGCGATGGC 2031
 QY 115 ACCTATCAGTGAATTTGCAAAACCAATCCGTCACTACCAATTTAAACCAACCAATCCAT 174
 DB 2030 AAGTATGAGTGGGTGACCAAGCCAAAGTCCGTGATCAATTTGCCACTCAGGCAAG 1971
 QY 175 GTTCAAAATTTGGGGTATAGTCTTGTGGTGGGTGAAACAAAGGCTTACCTCACC 234

Db 1970 GTGCCCTAAGCTGGGAATGATGCTGTCGGTTGGGGGTGCAACAACGGGTCTCACTCAAG 1911
 Qy 235 GGTGGGTATATGCTTAACAGAGGACATTTTATGGGCTTACAAAGA -CAAGTTCAACA 293
 Db 1910 GGTGGGTATCTCGCCCAACAAGAGGTATCTCATGGGTGACCAAGATGGCGTTGACGA 1851
 Qy 294 AACCAATTAATTTGGCTCCCTCAACCAAGGCTCAGCTATTTGAGTTGATCTCTCC - 349
 Db 1850 CGCCATTAATCTTGATCTCCCTCACTCAAGCTCACTTGGCCGGTGGCTCTTTTCAAG 1791
 Qy 350 ----AGGAGAGGAAATCTATGCCCCCATTCAGAGTCTGCTTCAATGGTTAATCTGAC 405
 Db 1790 GTGGAAGGAAGATTCTATGTAACCTTTTAAAGACATTTTAAACGATGATCAACCCCAAC 1731
 Qy 406 GACATTTGTTTGGGGGATGGGATATACGAAATGAACTGGCTGATGCTATGCTGCAAG 465
 Db 1730 GATCTGATGATGGCGGTGGGATATCTCAACATTAATGCTGCGATGCAATGATGATCT 1671
 Qy 466 GCMAAGGTGTTTGAATGATGCTGATTTGCAAGAGAGTGGAGGCTTTCATGGAATCTATGTT 525
 Db 1670 GCCCGGTCTGATGATGATGATCTGATGATGATGATGATGATGATGATGATGATGAT 1611
 Qy 526 CCATCTCCCGGAATCTAGACCCCGAATTTTATGCTGCTGCAACCAAGAGAGGCTGCAAC 585
 Db 1610 CCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1551
 Qy 586 AACGTGATTAAGGCAACAAGCAAGCAAGTTCAGCAATCATCAAAACATCAAGGCG 645
 Db 1550 AATCTATCAAGGTTTCCAAAGAGGCAATGAGCAAGGATTAATCCAGACATACCGAAC 1491
 Qy 646 TTTAAGGAAGCAACCAAGTGAACAAGGTGTTGCTGCTGAGCTGCCAACAAGAGAG 705
 Db 1490 TTCAAGGCGCAACCAAGGTTGACCAAGGTATGATGATGATGATGATGATGATGATGAT 1431
 Qy 706 TATAGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 764
 Db 1430 TACAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1371
 Qy 765 CAGAAATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 824
 Db 1370 GAAGGAGAAATGAAATCTCCCGCTGACCTTGTATGCTGCTGCTGATCAGCAGCA 1311
 Qy 825 TGTTCCTTTTCAATTAATGAGGCTCAGAAACATTTGTATGACAGGCTGATGATGATGAT 884
 Db 1310 GGTTCCTTTTCAATGAGGCTCAGAAACATTTGTATGATGATGATGATGATGATGATGAT 1251
 Qy 885 CATGCGGAGAACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 944
 Db 1250 TGTGAAAGAACTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1191
 Qy 945 ATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004
 Db 1190 ATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131
 Qy 1005 CAACCATCTGGAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1064
 Db 1130 CAACCATTTGGGCAACAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1071
 Qy 1065 GGAATTTCTCAAGAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1124
 Db 1070 GGAGATCTCAAGAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1011
 Qy 1125 GCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1184
 Db 1010 GCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
 Qy 1185 CAGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1244
 Db 950 CAGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 891
 Qy 1245 TTGCAACAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1304
 Db 890 CATGCAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 831

Qy 1305 TCTTGCTGAGCTGAGACATGAAATCCAGTTTAAAGCTGAAATAGAGGAAATTCACATC 1364
 Db 830 GCTGGCTGAGCTCTGACAAAGAAATCCAAATTCAGAGAAAGGACAGAGAAAGTTCCATTC 771
 Qy 1365 ATTCCACCCAGTTGGCTTACCATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1424
 Db 770 TTTCACCCCGGTTGCGTATCTCTGAGCTTACTTCAACCAAGGCTCTTGTGATGCCCAAG 711
 Qy 1425 TACACCAAGTGTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1484
 Db 710 TACCCAGAGTGTAAAGCGCTTGGCTAAGAGAGGCGGATGTAGAGAAATATCATGCGGCC 651
 Qy 1485 TTGTGTTGATTTGGCCCCCAGAGAAATTAATTAATTTCTGAGTCAAGT 1531
 Db 650 GTGCAATGGGTGTCTCCCGCAATTAACATGATGCTAAGATCAAGT 604

RESULT 5
 CK277950
 LOCUS 988 bp mRNA linear EST 03-AUG-2004
 DEFINITION EST724028 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POAB302 5' end, mRNA sequence.
 ACCESSION CK277950
 VERSION CK277950.1 GI:39834928
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 988)
 Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
 Generation of ESTs from abiotic stressed potato tissue
 Unpublished (2003)
 Other ESTs: EST724029
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@igr.org
 Clones can be requested from the University of Arizona Genomics
 Institute via http://genome.arizona.edu/orders/
 Seq primer: ATT TAG CTC ACA CTA TAG.
 FEATURES
 source
 1..988
 location/Qualifiers
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POAB302"
 /tissue_type="abiotic stress treated leaf and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="potato abiotic stress cDNA library"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: Solanum tuberosum var. Kennebec plants were
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation of
 the soil with 150 mM NaCl and tissues were harvested at
 following application of the salt stress (leaves: 2hr,
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
 Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering
 applications. Drought stressed plants were harvested after
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
 and 5d). Set 3 were grown under the standard conditions
 and then were cold stressed by placement at 4 C. Cold
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d. Set 4 were grown under the standard conditions and
 then were heat stressed by placement at 35 C. Heat
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d and 4d and heat-stressed roots were harvested at 6 hr,

12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

Query Match 45.6%; Score 698.8; DB 7; Length 988;

Best Local Similarity 82.3%; Pred. No. 5.5e-200; Mismatches 172; Indels 0; Gaps 0;

Matches 802; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

177 TCCTCAAAATGAGGGGATGATCTGTGGGTTGGGGTGAACAACAGGCTACCCCTGACCGG 236
 15 TTCCTCAATTAAGGGGTTATGCTTGTGTGATGGGGTGAACAACATGTTGACGCTTGA 74
 237 TGGTGTATGCTTAACAGAGAGACATTTATGCGCTCAACAAAGACAGATTCACCAAGC 296
 75 AGGTGTATGCTTAACAGAGAGATTTATGCGCTCAACAGATTCACCAAGC 134
 297 CAATTACTTGGGCTCCCTCAACCAAGGCTGACCTATGAGTTGATGATCTCTCCAGAGGA 356
 135 CAATTACTTGGGCTCTTACACAGGCTCTACTATTCAGTTGGTTCCTTAATGAGGA 194
 357 GGAATCTATGCGCCCATTCAGAGTCTGCTCAATGTTAATCTGACGATTTGTT 416
 135 GGAATCTATGCGCCCTTAAAGCATATTTCCATGTCATTCAGATGATGATGTT 254
 417 TGGGGGATGAGATATCAGACATGAACTGCTGATGTCATGCGCAAGGCAAGGTT 476
 255 TGGAGGATGAGCATCAGACATGAAATTTGGCAGATGTCATGCAAGGCTAAAGTTT 314
 477 TGAATCTGATTTGCAAGACATGAGGCTTCAATGATGATGATGTTCCCTCCCGG 536
 315 TGAATCTGATTTCAAAAGACGCTGAGGCTCAATGATGATGATGATGATGATGATG 374
 537 AATCTAGCAGCCGATTTGATGCTGTCGCAACCAAGAGAGGCTGCAACAGATTTAA 596
 375 TATCTATGACCTGATTTATGCTGCAACCAAGGCTCACTGCTAACAATGTCATTTA 434
 597 GGGCAAAAGCAAGCAAGTTTCAAGCAATCAAAAGCATCAAGGCTTTAAGGAAAC 656
 435 AGGAACCAAGAAAGCAAAATGATCAATCAATTAAGATTAAGGAGTTTAAAGAGA 494
 657 CACCAAGTGAACAAGTGTGCTCTGAGCTGCAACAGAGAGTATAGCAATTT 716
 495 CACCAAGTGAACAAGTGTGCTCTGAGCTGCAACAGAGTATAGCAATTT 554
 717 GGTGTAGGCTTAAATGACACATGAGATCTCTGCTGCTGAGCAGAAATGAGGC 776
 555 GGTGTAGGCTTAAATGACACATGAGATCTCTGCTGCTGAGCAGAAATGAGGC 614
 777 TGAATTTCTCTTCCACTTGTATGCAATGCTGCTGCTGATGAGAAATGTTCTTTCAT 836
 615 TGAATTTCTCTTCCACTTGTATGCAATGCTGCTGCTGATGAGAAATGTTCTTTCAT 674
 837 TGAATTTCTCTTCCACTTGTATGCAATGCTGCTGCTGATGAGAAATGTTCTTTCAT 896
 675 CAATGGAACCCCAACCAACATTTGTCCAGGCTCATGATTTGGCCATTAAGAGGA 734
 897 CACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 956
 735 CACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 794
 957 TGAATTTCTCTTCCACTTGTATGCAATGCTGCTGCTGATGAGAAATGTTCTTTCAT 1016
 795 TGAATTTCTCTTCCACTTGTATGCAATGCTGCTGCTGATGAGAAATGTTCTTTCAT 854
 1017 AATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076
 855 TGAATTTCTCTTCCACTTGTATGCAATGCTGCTGCTGATGAGAAATGTTCTTTCAT 914
 1077 GAGCAACGTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1136
 915 AAGCAATGTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 974

Query 1137 TCCGACCATGTTG 1150
 Db 975 CCTGACCATGTTG 988

RESULT 6

CK275352

LOCUS

DEFINITION

EST721430 potato

clone FOADM85 5' end, mRNA sequence.

CK275352

CK275352

KEYWORDS

SOURCE

ORGANISM

Solanum tuberosum (potato)

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamiales; Solanales; Solanaceae; Solanum.

1 (bases 1 to 914)

Buell, C.R., Hart, A., Zismann, V., Karamecheva, S.A. and Baker, B.

Generation of ESTs from abiotic stressed potato tissue

Unpublished (2003)

CONTACT: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@igr.org

Clones can be requested from the University of Arizona Genomics

Institute via http://genome.arizona.edu/orders/

Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

1..914

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="FOADM85"

/tissue_type="abiotic stress treated leaf and root tissue"

/lab_host="DH10B-Tona"

/clone_id="potato abiotic stress cDNA library"

/note="Vector: PCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

supplier: Solanum tuberosum var. Kennebec plants were

grown from cuttings on a 16hr light/8 hr dark cycle at 25

C for 3-4 weeks. Abiotic stress conditions were applied to

four separate sets of plants. Set 1 involved saturation of

the soil with 150 mM NaCl and tissues were harvested at

following application of the salt stress (leaves: 2hr,

6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).

Set 2 were grown under the standard conditions and then

were water stressed by withdrawal of further watering

applications. Drought stressed plants were harvested after

cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d

and 5d). Set 3 were grown under the standard conditions

and then were cold stressed by placement at 4 C. Cold

stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,

and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,

2d. Set 4 were grown under the standard conditions and

then were heat stressed by placement at 35 C. Heat

stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,

2d and 4d and heat-stressed roots were harvested at 6 hr,

12 hr, 1 d, and 4d. RNA was isolated from all tissues and

equal RNA from each tissue and stress was pooled to

construct the cDNA library. RNA sample."

ORIGIN

Query Match 42.5%; Score 651.4; DB 7; Length 914;

Best Local Similarity 83.5%; Pred. No. 1.3e-185;

Matches 739; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

649 AAGAAAGCACCACCAAGTGAAGGTTGTTCTGTGACTGCAACACAGAGAGTAT 708
 1 AAGAAAGCACCACCAAGTGAAGGTTGTTCTGTGACTGCAACACAGAGAGTAT 60
 709 AGCAATTTGTTGAGGCTTAATGACACCATGAGAAATCTTTGCTGCTGTGAGAGA 768

Db 61 AGACAGTGGCTGTGGCTTATGATACATGAGAAACCTCTTCTCTTGTGATGAGA 120
Qy 765 AATGAGGCTGAGATTTCTCTTCCACCTTGTATGCCATTCCTGTGTATGAGAAATGTT 828
Db 121 AATGAGGCGGAAATCTCTTCCACCTTGTATGATGCTGTATGCTTGTATGCTTGTG 180
Qy 829 CCTTCATTATGAGAGCCCTGAGAACATTTTGTACAGGGGTGATGATCTGTGCATC 888
Db 181 CCTTCATTATGAGAGCCCAAAACATTTTGTCCAGGGCTCATTTGATTTGGCCATC 240
Qy 889 GCGAGAGAACATTTGATGGTGGAGATGATCTTCCAGAGTGTGACACCAAAATGAATCT 948
Db 241 AAGAGAGAACATTTGATGGTGGAGATGATCTTCCAGAGTGTGACACCAAGATGAGTCA 300
Qy 949 GTGTTGTGATTTTCTTGTGGGGCTGTATTAAGCCCAATCTATAGTTAGTTACAA 1008
Db 301 GTGCTGTGATTTCTTGTGGAGCTGTATTAAGCCCAATCTATAGTTAGTTAGTCAAC 360
Qy 1009 CATCTGGGAAACATGATGGTATGATCTTGGGCTCCCAAACTTCGCTCCAGAGAA 1068
Db 361 CATTTGGGTAACATATGAGAAATGATCTGTGGGGCTTCAGACTTCCGCTCAAGAGAG 420
Qy 1069 ATCTCCAGAGAGAACCTTGTGACGATTAAGTTCAGAGAGATGCCATCTCTATAGAGCT 1128
Db 421 ATCTGGAAAGAGAACTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 480
Qy 1129 GGTGAACATCCGACCATGTTGTTTATTAAGTATGCTTATGCTTATGAGAGCT 1188
Db 481 GGAAGAGACCTGACCATGTTGTTGATTAAGTATGATGATGATGATGATGATGATG 540
Qy 1189 AGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1248
Db 541 AGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Qy 1249 CACAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1308
Db 601 CACAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Qy 1309 GGTGAGCTGAGCATTAAGATTCAGATTTAAAGCTGAAATGAGGAAATTCACATCT 1368
Db 661 GGTGAGCTGAGCATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 720
Qy 1369 CACCAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1428
Db 721 CATCCGAGGAGCAATCTCAAGCTATCTTACCAAGGCTCTCTGTGATGATGATGATG 780
Qy 1429 CCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1488
Db 781 CCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Qy 1489 GTTGAATGGCCCGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1533
Db 841 GTTGAATGGCCCGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 885

RESULT 7
LOCUS CK279064 936 bp mRNA linear EST 03-AUG-2004
DEFINITION EST125142 potato abiotic stress cDNA library Solanum tuberosum cDNA
clone POAE989 5' end, mRNA sequence.
ACCESSION CK279064
VERSION CK279064.1 GI:39836042
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
AUTHORS Generation of ESTs from abiotic stressed potato tissue
JOURNAL Unpublished (2003)

COMMENT Other ESTs: EST725143
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igrr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATG TAG GTG ACA CTA TGG.
Location/Qualifiers
1. 936
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAE989"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="potato abiotic stress cDNA library"
/note="Vector: PCW5Sport6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

FEATURES
source

Query Match 42.3%; Score 648.4; DB 7; Length 936;
Best Local Similarity 81.5%; Pred. No. 1.1e-184;
Matches 763; Conservative 0; Mismatches 171; Indels 2; Gaps 1;

Qy 485 ATTGCAAGAGAGTGGAGGCTTACATGAGATTCATGCTCCCGGAATCTAG 544
Db 1 ATCTGCAAAAGAGAGCTGAGGCTTACATGAGATTCATGCTCTTCTGTGATCTATG 60
Qy 545 ACCCGAATTCATGCTGCAACCAAGAGAGCGTCCCAACAGCTGATTAAGGACAA 604
Db 61 ACCCTGACTTCAATGACGTAACCAAGAGAGCGTCCCAACAGCTGATTAAGGACAA 120
Qy 605 AGCAAGAGCAAGTTCAGCAATATCATCAAGAGATCAAGCGTTTAAAGAGCCACCAAG 664
Db 121 AGAAAGAACAGATTATCAAAATGTTAAAGATTAAGGAGTTCAAGAGAGAAACAAAG 180
Qy 665 TGCAGAGAGGTGTGCTGCTGAGTCCCAACAGAGAGATGATGCAATTTGGTTG 724
Db 181 TGCAGAGATGATGTTCTATGAGCTGCCAACAGAGAGATGATGATGATGATG 240
Qy 725 GCTTATGACACCATGAGAAATCTTGTGCTGTGTGACAGAAATGAGGCTGAGATT 784
Db 241 GCTTATGACATCATGAGAAACCTTTAGCTGTGTGATGAGAAATGAGGCTGAGATT 300
Qy 785 CTCCTTCAACCTTGTATGATGCTGTGTGATGAGAAATGTTCTTTTCAATTAAGAA 844
Db 301 CTCCTTCAACCTTGTATGATGCTGTGTGATGAGAAATGTTCTTTTCAATTAAGAA 360
Qy 845 GGCCTCAGAACCTTTGATGAGAGGCTGATGATCTTGCATGCGAGAGAACCTTTGA 904

Db 361 GCCCTCAAAACACTTTTGTCCAGGTCTTATGATTTGGCCATAAGAAACACTTTAA 420
QY 905 TTGTGAGATGATCTTCAAGAGTGGTTCAGACCAAAATGAAATCTGTGTGTTGATTTTC 964
Db 421 TTGTGAGATGATCTTCAAGAGTGGTTCAGACCAAAATGAAATCTGTGTGTTGATTTTC 480
QY 965 TTGTGAGAGTGGTTCAGAGCAACATCTTATGTTTACCAACATCTGGGAAACAATG 1024
Db 481 TTGTGAGAGTGGTTCAGAGCAACATCTTATGTTTACCAACATCTGGGTAACAATG 540
QY 1025 ATGTGATGATCTCTCGGTCCAAACCTTCGCTCCAGGAATCTCCAGACCAACG 1084
Db 541 ACGGATGATCTCTCTCGGTCCAAACCTTCGCTCCAGGAATCTCCAGGAATGATG 600
QY 1085 TTGTGAGATGATCTTCAAGAGTGGTTCAGACCAAAATGAAATCTGTGTGTTGATTTTC 1144
Db 601 TTGTGATGATCTGTGTGTTGATGAACTCTATCTATGATCTGTGAGACCTTGACC 660
QY 1145 ATGTGTTGTTATTAAGTATGCTTACGTAAGGAGATAGCAAGAGCCATGATGAGT 1204
Db 661 ATGTGTTGTTATTAAGTATGCTTACGTAAGGAGATAGCAAGAGCCATGATGAGT 720
QY 1205 ACATCTCAGAGTATTCATGAGTGGTGAAGAGACCAATGTTTGAACAACATGAGAG 1264
Db 721 ACATCTCAGAGTATTCATGAGGCGGAGAGAGCAACATGTTATGACCAACACTGTGTAG 780
QY 1265 ATTCCTTTTACCTCTCTATTAATCTTGAAGTGGTCTCTCTCTGAGTGAACACTA 1324
Db 781 ACTCTCTTTTGGACCTCCAAATTAATCTTGAAGTGGTCTCTCTCTGAGTGAACACTA 840
QY 1335 GAATCCAGTTTAAAGTGAATAAG--AGGAAATTTCCACTATTCACCCAGTTGCTAC 1382
Db 841 GCATTCAGTCTCAAGCTGAAAGGAGGAGGAGGAGGAGTCCACTCTTCCACTGTGTCTAC 900
QY 1383 CATTCAGCTATCTGACCAAGGCTCTCTGTGTTCC 1418
Db 901 TATCTCAGCTACCTCACCAGGCTCCCTGTGATCC 936

RESULT 8
LOCUS CO085839 865 bp mRNA linear EST 16-JUN-2004
DEFINITION GR_Ea03E19.f GR_Ea Gossypium raimondii cDNA clone GR_Ea03E19 5',
mRNA sequence.
ACCESSION CO085839
VERSION CO085839.1 GI:48776473
KEYWORDS EST
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 865)
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 03 row: E column: 19.
FEATURES
source 1..865
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea03E19"
/issue_type="whole seedlings"

ORIGIN
Query Match 41.5%; Score 636.2; DB 7; Length 865;
Best Local Similarity 83.5%; Pred. No. 5.2e-181;
Matches 722; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
/dev_stage="first true leaves"
/lab_host="DHI0B"
/clone_id="GR_Ea"
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."
QY 426 GGATATGACCAACATGAACCTGCTGATGCCATGCGCCAGGGCAAGTGTGATGATGCA 485
Db 1 GACATATGATGACATGAACCTAGCTGATCTATGCTGAGGCGCAAGTGTTCGACATGCA 60
QY 486 TTGTGAGAGAGATGAGAGCTTACATGATGATGATGATGATGATGATGATGATGATGAT 545
Db 61 TCTGAAAGAGAGATGAGAGCTTACATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 546 CCGGATTTCAATGCTGCAACCAAGAGAGAGCTGCAACAGTATTAAGGAGGACAAA 605
Db 121 TCTGATTTCAATGCTGCTTACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 180
QY 606 GCAAGAGCAAGTTCAGCAATTCATCAAGATCAAGAGCGTTTAAAGAACCACTAAAGT 665
Db 181 GAAAGAACAGATTCAGCAGTATCAAGAGATCAAGAGATTCAGAGAGAGAGAGAGAGAGT 240
QY 666 GGACAGAGAGTGTGCTGCTGAGAGCTGCAAGAGAGAGATGATGATGATGATGATGATGAT 725
Db 241 GGACAGAGTGTGCTGCTGAGAGCTGCAAGAGAGAGATGATGATGATGATGATGATGATGAT 300
QY 726 CCTTATGACACCAATGAGAAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 785
Db 301 GCTAATGACACCGTGAAGAGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 360
QY 786 TCTTTCACCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 845
Db 361 TCTTTCACCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 846 CCTTACAGAACCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 905
Db 421 CCACAAAACACTTTGTTTCCAGGGTATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 906 TTGTGAGATGATCTTCAAGAGTGGTTCAGACCAAAATGAAATCTGTGTGTTGATTTTCT 965
Db 481 TGGAGAGATGATCTTCAAGAGTGGTTCAGACCAAAATGAAATCTGTGTGTTGATTTTCT 540
QY 966 TGTGGGGGCTGGTATCAAGAGTGGTTCAGACCAAAATGAAATCTGTGTGTTGATTTTCT 1025
Db 541 TGTGGGGGCTGGGATCAAGAGTGGTTCAGACCAAAATGAAATCTGTGTGTTGATTTTCT 600
QY 1026 TGTGATGATCTCTCGGTCCAAACCTTCGCTCCAGGAATCTCCAGAGCAACG 1085
Db 601 TGGCATGATCTGTAGAGACCCCAACCTTCGCTCCAGGAATCTCCAGAGCAACG 660
QY 1086 TGTGACATATGATCAAGAGTGGTTCAGACCAAAATGAAATCTGTGTGTTGATTTTCT 1145
Db 661 TGTGATGATCTGTGTGTTGATGAGCAATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 1146 TGTGTTGTTATTAAGTATGCTTACGTAAGGAGATAGCAAGAGCCATGATGAGT 1205
Db 721 TGTGTTGTTATTAAGTATGCTTACGTAAGGAGATAGCAAGAGCCATGATGAGT 780
QY 1206 CACTTCAGAGATATTCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1265
Db 781 CACATCAGAGATATTCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 1266 TTCCCTTTAGCTGCTCTATATC 1290
Db 841 TTCCCTGTTGGCTGCCCTCCATATC 865

RESULT 9
CK320178 906 bp mRNA linear EST 11-MAY-2004

LOCUS L2P08a02 Populus stem seasonal library Populus deltoides cDNA, mRNA
DEFINITION sequence.
CK320178
ACCESSION CK320178 GI:47108601
VERSION
KEYWORDS
SOURCE
ORGANISM
Populus deltoides
Populus deltoides
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malviales; Salicaceae; Saliceae; Populus.
1 (bases 1 to 906)
Park, S. and Han, K.-H.
Gene expression profile during seasonal growth cycle in poplar tree
Unpublished (2003)
JOURNAL
AUTHORS
TITLE
COMMENT
Department of Forestry
Kyoung-Hwan Han
Michigan State University
126 Natural Resources, East Lansing, MI 48824-1222, USA
Tel: 517 353 4751
Fax: 517 432 1143
Email: hanky@msu.edu.

FEATURES
source
1..906
/organism="Populus deltoides"
/mol_type="mRNA"
/strain="IL-129"
/db_xref="taxon:3696"
/tissue_type="stem"
/dev_stage="1 year old"
/clone_lib="Populus stem seasonal library"

ORIGIN

Query Match 41.3%; Score 633; DB 7; Length 906;
Best Local Similarity 82.4%; Pred. No. 56-180;
Matches 726; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

416 TTGGGGGATGGATATCAGCAACATGAACTGGCTGATCCATGGCCAGGCAAGGTT 475
26 TTACGGCCGGGTCATTAAGTGAATGAATCTGGCAGATGCATGGCGGCAAGGTTT 85
476 TTGACATCATTTTGCAGAACGCTTGAGGCTTACATGGAATCATGTTCCATCCCG 535
86 TCGACATTCATCTGCAAAAGCAATGAGGCTCTACATGGAATCATGTAACCATCCCTG 145
536 GAATCAGACCGGATTCATTTGCTGCGCAACAGAGAGGCTGCAACATCGATTA 595
146 GAATTTACATCTCTGATTCATTTGCTGCGCAACAGGCTCAGCTGTAACATGTTATCA 205
596 AGGGCAAAAGCAAGAGCAAGTTCAAGCAATCATCAAGACATCAAGGCTTTAAGAG 655
206 AAGGCACTTAAGAAAGCAAGTCAGCAAAATTAAGCAATTAAGGAGTTCAAGAGA 265
656 CCACCAAGTGAACAAGTGTGTTCTCTGTGAATCCCAACAAGAGAGTATTAACAAT 715
266 AAAACAAGTGAACAAGTGTGTTGTGTGAATCCCAACAAGAGAGTACAGTAATA 325
716 TGGTGTGGGCTTAATGACACATGAGAAATCTCTGGCTCTGTGACAGAAATGAGG 775
326 TTGTTGTGGGCTGAATGACACATGAGAAATCTCTGGCTCTGTGAGGAAGAGTAAT 385
776 CTGAGATTTCTCTTCAACCTTGATGACATGCTCTGTGATGAGAAATGTTCTTTCA 835
386 CAGAGATATCCCATCAACCTTGATGCTCTGTGATGATTAAGAAATATTCCTTTCA 445
836 TTAATGAAAGCCCTCAAGAACTTTTGTACCAAGGCTGATTAATCTTGCATCGGAGA 895
446 TAAATGGAAGCCCAAGAACTTTTGTTCAGGACTTGTGATTTGGCTATTAAGAGA 505
896 ACACTTTATTTGTGAGATGACTTCAAGAGTGTCAAGACCAAAATGAATCTGTGTGG 955

Db 506 AAGTTGATTTGTGAGATGACTTTAAGATGTCAGCCAAATGAATCTGTTCTGG 565

Qy 956 TTGATTTTCTTGTGGGGCTGTATCAAGCAACATCTATAGTTAGTCAACATCTGG 1015

Db 566 TTGATTTCTTGTGGGGCTGTATTAAGCCAACTCAATAGTGAAGTAAACCATCTGG 625

Qy 1016 GAAACAATGATGATGATATCTCTGCTCCCAAACTTCTCGCTCCAAAGAAATCTCA 1075

Db 626 GAAACAATGATGATGATATCTCTGCTCCCAAACTTCTCGCTCCAAAGAAATCTCA 685

Qy 1076 AGAGCAAGTGTGATGATATGTCACAGCAAGATGCCATCTCTATGAGCTGTGTAAC 1135

Db 686 AGAGCAAGTGTGATGATATGTCACAGCAAGATGCCATCTCTATGAGCTGTGTAAC 745

Qy 1136 ATCCGACCATGTTGTTGATTAAGTATGCTTACGATGAGGAGTAAAGAGACCA 1195

Db 746 ATCTGATCATATGTTAGTAATCAAGTATGTCATATGTTGAGATGACAGAGACTTA 805

Qy 1196 TTGATGATGATCATCTCAAGATATTCATGAGTGGGAAAGAACCATTTGTTGCAACA 1255

Db 806 TGGATGAGTACACCTCGAGATATTCATGAGTGGGAAAGAACCATTTGTTGCAACA 865

Qy 1256 CATGTGAGATTCCTTTTACGCTGCTCTATATCTTGGAC 1296

Db 866 CCGTGAAGACTCCCTCTGCTGACCAACATCATCTCGAC 906

RESULT 10
COL13870 898 bp mRNA linear EST 16-JUN-2004
LOCUS COL13870
DEFINITION GR_Eb014121.r GR_Eb Gossypium raimondii cDNA clone GR_Eb014121
3', mRNA sequence.
COL13870
ACCESSION COL13870 GI:48812557
VERSION
KEYWORDS
SOURCE
ORGANISM
Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 898)
Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and
Wing, R. A.
Global assembly of Cotton ESTs
Unpublished (2004)
JOURNAL
COMMENT
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 014 row: I column: 21.

FEATURES
source
1..898
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb014121"
/tissue_type="Floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Clones
plated/picked by Agt. More glycerol clones held in -80."

ORIGIN

Query Match 41.0%; Score 628.2; DB 7; Length 898;
Best Local Similarity 81.2%; Pred. No. 1.4e-178;

Matches	729;	Conservative	0;	Mismatches	169;	Indels	0;	Gaps	0;
QY	145	GTCACTACCAATTTAAACCAACCCCAATGTTCCAAATTTGGGGGTGATGCTGGGT	204						
Db	1	GTCAATATGAATTCAGACTGATACCATGCCCTTAATTTGGGGGTGATGCTGGGT	60						
QY	205	TGGGGTGAACACGCTCTACCTCCACCGGTGTTATTTCTAAGAGAGACATT	264						
Db	61	TGGGGAGAAACATGTTCAACCTCACCGGTGTTATTTAGCTAACAGAGGTATC	120						
QY	265	TGATGGGCTACAAAGACAAAGATTCACCAAGCCATTTCTGGCTCCTCACCCAGCC	324						
Db	121	TCTTGGGCTACAAAGACAAAGGTCTAATTTCTTTGGTTCATTTGATCAAGA	180						
QY	325	TGAGCTATTCAGATTGATCTCTCCAGAGAGAGAAATCTATGCCCATTTCAAGAGTCTG	384						
Db	181	TCAACGATCCGAATTTGGTCTTACAAATGAGAGAAATTTATGCTCATTTAAGAGTCTT	240						
QY	385	CTTCCATATGTTAATCTTACGACATTTGTTGGGGATGGGATATCAGCAATGAAAC	444						
Db	241	CTTCCATATGTTAATCTTACGACATTTGTTGGGGATGGGATATGATGACATGAAAC	300						
QY	445	CTGGCTGATGCCATGGCCAGGCGCAAGGTGTTGACATGATTTGAGAGAGAGATTGAGG	504						
Db	301	CTAGCTGATGCAATGGCTAGGGCCAGGTTTTCGACATGATCTGCAAAAGCAACTGAGA	360						
QY	505	CCTTACATGGAATTCATGTTCCATCCCGGATCTACGACCCGATTTCAATGCTGCGC	564						
Db	361	CCCTACATGGAATTCATGTTCCATCCCGGATCTACGATCTGATTTCAATGCTGCGC	420						
QY	565	AAACCAAGAGAGCGTGCACCAACGATTTAAGGCAACAGCAAGAGCAAGTTACGAA	624						
Db	421	AAACCAAGAGAGCGTGCACCAATATGTCATCAAGGGGACCAAGAAAGCAAGTTACGAG	480						
QY	625	ATCATCAAGACATCAAGGCTTTAAGAGAGCCACCAAGTGAAGAGTGTGCTGCTG	684						
Db	481	GTATCAAAAGACATCAAGGCTTTAAGAGAGCCACCAAGTGAAGAGTGTGCTGCTG	540						
QY	685	TGAGCTGCCAAGACAGAGAGTATAGCAATTTGTTGAGGCTTTATGACACCAATGAG	744						
Db	541	TGAGCTGCCAAGACAGAGAGTATAGCAATTTGTTGAGGCTTTATGACACCAATGAG	600						
QY	745	AATCTCTGGCTGCTGTGACAGAAATGAGGCTGAGATTTCTCTTCCACTTGTATGCC	804						
Db	601	AACCTTATGCTCTCTTGGAGAAAGATGATCAAGAAATTTCTCTTCCACTTGTATGCT	660						
QY	805	ATTGCTGTGTGATGAGAAATGTTCTTTCAATTAAGAAAGCCCTCAGAAACCTTTGTA	864						
Db	661	ATTGCTGTGTGATGAGAAATGTTCTTTCAATTAAGAAAGCCCTCAGAAACCTTTGTT	720						
QY	865	CCAGGGCTGATGATCTTCCATCGGAGAGAAACATTTGATTTGGTGAATGACCTTCAAG	924						
Db	721	CCAGGGCTGATGATTTGGCTATTCAAAGAACTGCTATTTGGAGAGATGATCTTCAAG	780						
QY	925	AGTGTGACAGCAAAATGAATCTGTGTTGTTGATTTCTTGTGGGGCTGGTATCAAG	984						
Db	781	AGTGTGACAGCAAAATGAATCTGTGTTGTTGATTTCTTGTGGGGCTGGTATCAAG	840						
QY	985	CCAACTCTATATGTTATGTAACCAACATCTGGGAGAAACAAATGATGATGATCTCTGG	1042						
Db	841	CCAACTCTATATGTTATGTAACCAACATCTGGGAGAAATGATGATGATGATCTCTGG	898						

RESULT 11

LOCUS BG044525 670 bp mRNA linear EST 23-JUL-2004
 DEFINITION saae9e07.y1 Gm-cl059 Glycine soja cDNA clone GENOME SYSTEMS CLONE
 ID: Gm-cl059-589 5' similar to TR:095SV4 095SV4
 MYO-INOSITOL-1-PHOSPHATE SYNTHASE. ; mRNA sequence.
 ACCESSION BG044525 GI:12491342
 VERSION BG044525.1
 KEYWORDS EST.
 SOURCE Glycine soja

ORGANISM	Glycine soja
REFERENCE	1 (bases 1 to 670)
AUTHORS	Shoemaker, R., Kaim, P., Vodkin, L., Espelting, J., Coryell, V., Khanna, A., Bolla, B., Maitra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Thelting, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE	Public Soybean EST Project
JOURNAL	Unpublished (1999)
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu
FEATURES	When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
source	High quality sequence stop: 422. Location/Qualifiers 1..670 /organism="Glycine soja" /mol_type="mRNA" /db_xref="taxon:3848" /clone="GENOME SYSTEMS CLONE ID: Gm-cl059-589" /tissue_type="whole seedling, 2 week old, etiolated" /lab_host="DH10B" /clone_id="Gm-cl059" /note="Vector: pBlueScript II SK+, Site 1: EcoRI, Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 2 week old etiolated whole seedlings of P1468916. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBlueScript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed in the laboratory of Dr. Randy Shoemaker at Iowa state university."
ORIGIN	
Query Match	40.8%; Score 625.8; DB 4; Length 670;
Best Local Similarity	96.9%; Pred. No. 6.8e-178;
Matches	649; Conservative 0; Mismatches 18; Indels 3; Gaps 1;
QY	474 GTTTGACATCGATTTTGCAGAAAGCAATGAGGCTTACATGGAATTCATGCTCCCTCC
Db	1 GTTTGACATCGATTTTGCAGAAAGCAATGAGGCTTACATGGAATTCATGCTCCCTCC
QY	534 CGGAATCTACGACCGGATTTGATGCTGCGCAACCAAGAGAGCGTGCACCAACGAT
Db	61 CGGAATCTACGACCGGATTTGATGCTGCGCAACCAAGAGAGCGTGCACCAACGAT
QY	594 TAAGGGCAAAAGCAAGCAAGATTACCAATATCTCAAGACATCAAGGCGTTTAAAGA
Db	121 TAAGGGCAAAAGCAAGCAAGATTACCAATATCTCAAGACATCAAGGCGTTTAAAGA
QY	654 AGCCACCAAGTGAAGCAAGGTGTTGCTCTGTGACCTCCCAACAGAGAGATTAAGCA
Db	181 AGCCACCAAGTGAAGCAAGGTGTTGCTCTGTGACCTCCCAACAGAGAGATTAAGCA
QY	714 TTTGGTGTAGGCGCTTAATGACCATGAGAGATCTTGGCTGCTGAGACAGAAATGA
Db	241 TTTGGTGTAGGCGCTTAATGACCATGAGAGATCTTGGCTGCTGAGACAGAAATGA

QY	774	GGCGAGATTTCTCCCTTCACCTGTGTAGCCATTGCGTGTGATGAGAAAAATGTCCTT	833
Db	301	GGCGAGATTTCTCTTCACCTGTGTAGCCATTGCGTGTGATGAGAAAAATGTCCTT	360
QY	834	CATTATATGAAGCCCTCAGAACACTTTTGTACAGAGCGCTGATTTGATCTTGCATCGCAG	893
Db	361	CATTATATGAAGCCCTCAGAACACTTTTGTACAGAGCGCTGATTTGATCTTGCATCGCAG	420
QY	894	GAACACTTTGATTTGCTGTGAGATGACTTCAAGAGTGTCAAGACCAAAATGMAATCTGTGT	953
Db	421	GAACACTTTGATTTGCTGTGAGATGACTTCAAGAGTGTCAAGACCAAAATGMAATCTGTGT	480
QY	954	GGTGAATTTCTTGTGGGGCGGTGATCAAGCAACATCTATAGTTAGTTCAACATCT	1013
Db	481	GGTGAATTTCTTGTGGGGCGGTGATCAAGCAACATCTATAGTTAGTTCAACATCT	540
QY	1014	GGGAAACAAT--GATGATATGAATCTCTCGGCTTCACAAACTTCGCGCTCAAGGAAT	1070
Db	541	TGGGAAACATATGATGGGTATGAATCTCTCGGCTTCACAAACTTCGCGCTCAAGGAAT	600
QY	1071	CTCCAAGCAACGTTGTTGAAGATATGCTCAACAGCAATGCAATGTCATCTCTATGAGCTGG	1130
Db	601	CTCCAAGCAACGTTGTTGAAGATATGCTCAACAGCAATGCAATGTCATCTCTATGAGCTGG	660
QY	1131	TGAACATCCC 1140	
Db	661	TGAACATCCC 670	

RESULT 12	
C0072025	
LOCUS	
DEFINITION	C0072025 860 bp mRNA linear EST 15-JUN-2004
	GR_Ea30P03.r GR_Ea Gossypium raimondii cDNA clone GR_Ea30P03 3'
mRNA sequence.	

ACCESSION	CO072025
VERSION	CO072025.1
KEYWORDS	GI:48741506
SOURCE	EST.
	<i>Gossypium raimondii</i>

ORGANISM
Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosid II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 860)

REFERENCE

AUTHORS	Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., Udell, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and Wing, R. A.
TITLE	Global assembly of Cotton ESTs
JOURNAL	Unpublished (2004)
COMMENT	Contact: Rod A. Wing

Arizona Genomics Institute
The University of Arizona
Forbes Building, Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: <http://genome.arizona.edu>
Plate: 30 row: P column: 03.

FEATURES	Location/Qualifiers
SOURCE	1. 860

```

/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea30p03"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"

```

```

/clone_11b="GR_Ea"
/note="vector: pCMV_SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wentle Lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

```

ORIGIN	Query Match
	40.7%; Score 624.4; DB 7; Length 860;

Best Local Similarity 82.9%; Pred. No. 2e-177;
Matches 712; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

527 CACTCCCCGGAATCTACGACCCCGATTTCATTGCTGCCAACCAAGAGAGCGTGCCAACA 586

1 CACTCCCTGGATCTACGATCCGATCTCATTTGCTGCTAACCAAGTGACCGCCAAATA 60

587 ACGTGAATTAAGGGCACAAAGCAAGCAAGTTACGAAATATCTAAAGACATCAAGGCGT 646

61 ATGTCAATCAAGGGGACCAAGAAACAAGTTCAAGCAGTTCATCAAAAGACATCAAGGAGT 120

647 TTAAGCAAGCCACCAAAATGGACAGTGSTGTCTCTGTGGACTGCCAACAGAGAGST 706

121 TCAAGGAAAAA CAGGTGGACAGGTTGTGTACTCTGACATGCAACACCTGAGAGGT 180

181 A C A G C A A T G T C A T T C G T G G G G C T A A A T G A C A C C C T G A A A G C C T T A T G A G C C T T C T T T T G A G A 240
/0/ A T A G C A A A I T G T I G T A G G C C I T A A I G A C A C C A I G G A G A A I C I C I T T G C C T G C T G I G I G A C A /88

201 GAAATGAGGCGTGGAGATTTCTCTTCCACCGTTGGTATGCCATTGCGGCGGAAAAATG 826

241 AGAATGATCAGAGATTTCCTTCACACTTTGATGCTAATGCTTGTCGTTCTTGAAAATG 300

827 TTCCTTCATTATGGAAGCCCTCAGAACACTTTTGTACCAAGGGCTGATTGATCTTGCCA 886

301 TTCTTTCATCATGGCAGCCCAAAACACTTTGTTCCAGGTTGATTGATTGGCTA 360

887 TCGGAGGACACTTTGATTGGTGAGATGACTTCAAGAGTGTCAACCAAAATGAAT 946

361 TTCAAAGGACCTGCTCTGATTGGAGGAGATGACTTCAAGAGTGGCCAGACCAAGATGAAT 420

947 CTGTGTTGTTGATTTTCTTGTGGGGCTGTATCAAGCCACATCATATAGTTAGTTACA 1006

421 CTGTCTCGTGAATTTCTCTTGTGGGGCTGGGATCAAGCCAAACATGATAGTAGTTACA 480

1007 ACCATCTGGGAACATATGTAATCTCTGGCTCCACAACTTCCGCTCCAAAG 1066

481 ACCATCTGGGAAATTAATGATGGCATGAACTGTGAGCACCCCAAACTTCCGTTCCAAAGG 540

5A1 AGATCGGAGAGCCGCTGGTTCACGATATGGTGCAACAGCAAATGCCATCCTCATAGGC 600
106 / AAAATCTCCAAAGGACAAACGCTGTGTTGACGATATGGTGCAACAGCAAATGCCATCCTCATAGGC 1128

1127 CTGATGAACATCCCGACCATTGTCTTTAATAAGTAGTCCTTAAGTAAAGGGAGAATASCA 1186
341 AGATCCTCATTTAGAACCMATGTAATGTAATGCACAAGAAAATCCACTGATTAAGACC 000

601 CTGTGACACATCTTGATCATGTTGGTCATCAAGTATGTCACATATGTGGAGACAGCA 660

1187 AGAGAGCCATGATGAGTACCTTCAGAGATATTTCATGTGGTGGAAAGACACCATTTGTTT 1246

661 AGAGAGCATGATGAGTACACATCAGAGATATTCATGGAGGCAAGAACCATTTGTT 720

1247 TGCACAAACATGTGAGATTCCCTTTAGCTGCTCTTAATTATCTTGGACTTGGTCCCTTC 1306

721 TGCACACACATGTGAGATTCCCTGTTGGCTGCTCCCAATTATCCTAGATTGGTTC 780

1307 TTGCTGAGCTGAGCACTAGAATCCAGTTTAAAGCTGAAATGAGGAAATTCACATCAT 1366

781 TTGCTGAGCTTAGCACCAGATCCAGTTCAGAGGTGATGAGAGGCGCAAGTTCACATCTT 840

1367 TCACCCAGTTGCTACCAT 1385

841 TCCATCCTGTGGCTACAAT 859

RESULT 13
E211061

Accession	Length	Type	EST ID
CF211061	842 bp	mRNA	EST 01-AUG-2003
CA820007	11a	Fa D05 Cabernet Sauvignon flower bloom - CAB2 virus	

vinfiera_cDNA clone CAB2007_11a_Fa_D05 5', mRNA sequence.
 CF211061
 CCESSION

VERSION CP211061.1 GI:33405434
KEYWORDS EST.

SOURCE *Vitis vinifera*

ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
1 (bases 1 to 842)
Jones, K. and Cook, D.
Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
berries at various developmental stages
Unpublished (2003)

TITLE
JOURNAL Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcoc@ucdavis.edu
Seq primer: ACGGTACCGACATATGCC.
Location/Qualifiers

FEATURES
source
1..842
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultiivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CAB20007_Ira_Fa_D05"
/sex="Hermaphrodite"
/dev_stage="Bloom"
/clone_lib="Cabernet Sauvignon Flower bloom - CAB2"
/note="Organ: Flower - Bloom; Vector: pDNR; Site 1: S11;
Site 2: S11; CAB2 is a cDNA library of Vitis vinifera cv.
'Cabernet Sauvignon' Clone 8 berries. Samples were
collected at full bloom (80 to 100% flowers showing
dehiscence of calypters or caps and anthers fully
extended). Sampled vines were located at the University of
California, Davis, Experimental Vineyard. cDNAs were made
by oligo-dT priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGTATCAACGACGAGTGGCATTAAGCGCGG-3' and
5'-ATTGTAGAGCGGACGCGGCGACATG-3' (30)NN-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."

ORIGIN
Query Match 40.7%; Score 623.6; DB 7; Length 842;
Best Local Similarity 84.6%; Pred. No. 3.5e-177;
Matches 712; Conservative 0; Mismatches 129; Indels 1; Gaps 1;

QY 382 CTGCTTCAATGTTAATCTTACGACATTTGT- TTGGGGGATGGATATACCAACAT 440
DB 1 CTCCTCCCTATGTGAACCCAGATGACATTTGGATCTTGGGGATGGACATCAGTACAT 60
QY 441 GAACCTGGGTATGTCATGGCGGCAAGGTTTTCATGCATGATTTGGCAGAACAGAT 500
DB 61 GAACCTGGGTATGTCATGGCGGCTTAAGGTCCTGACATGATTTGCAACAAACAGTT 120
QY 501 GAGGCTTACATGATGATCCATGTTCACTCCCGAAATCTACGACCCGATTTTCATTC 560
DB 121 GAGGCTTACATGATGATCCATGTTCACTCCCGAAATCTACGACCCGATTTTCATTC 180
QY 561 TGCCCAACCAAGAGAGCGTGGCAACACGTGATTAAGGGCACAAGCAAGCAAGTTCA 620
DB 181 TGCCCAACCAATGCTCTCGGCAACACGTGATTAAGGGCACAAGCAAGCAAGTTCA 240
QY 621 GCAATTCATCAAGACATCAAGGCGTTTAAGGAACCAACCAAGGTCGCAAGGTTGT 680
DB 241 GCAATTCATTAAGATATTAGGAGTTTAAGGAGCAACCAAGGTCGCAAGGTTGT 300
QY 681 CCGTGTGACTGCCAACACAGAGGTATAGCAATTTGTTAGGCTTAAATGACACCAT 740
DB 301 GGTGTGACTGCCAACACAGAGCATAGTAAATTTGATTTGGGACTGAATGACACCAT 360
QY 741 GGAGATCTCTTGGCTGTGTGAGCAAGAAATAGGCTGATTTCTCTCCACCTTGA 800
DB 801

DB 361 AGAAATCTCTTGGCTTCTCTGATAGGAATGAGCTGAGATCTTCTCCCTCCACTTGA 420
QY 801 TGCCATTCCTGTGTGATAGGAATGTTCCCTTTCATTAATGGAAGCCCTCAGAACACTT 860
DB 421 TGCTCTGCTTGTGTGATAGGAATGTTCCCTTTCATTAATGGAAGCCCTCAGAACACTT 480
QY 861 TGTACCAAGGCTGATGATCTTGGCATTCGAGGAACAATTGATTTGGTGAATGACTT 920
DB 481 TGTTCAGAGGCTGATGATTTGGCTATTAAGAGGAACAAGTTGATTTGGTGGGATGATTT 540
QY 921 CAAGATGCTGACACCAAAATGCTGCTGTTGATTTCTTGGGGGCTGGAT 980
DB 541 TTAAGATGCTTAACCAAGATGAATCTGATTTGTTGATTTCTTGGGGGAGGAT 600
QY 981 CAAGCAACATCTATAGTATGTTAATCAACCATCTGGAAACATATGATGATATCTTC 1040
DB 601 CAAGCAACATCTATGTTGATGCTTAAACCATCTGGAAACATATGATGATATCTTC 660
QY 1041 GGCTTCAACAACTTCGCTTCAAGAAATCTCCAAAGCAACCTTTGACATATGCT 1100
DB 661 TGCTCTCAACCTTCGCTTCAAGAAATCTCCAAAGCAACCTTTGACATATGCT 720
QY 1101 CAACGCAATGCTATCTCTTATGAGCTTGTGACATCTCCGACATCCGACATGTTGTTATTA 1160
DB 721 TTCCAGCAACGCTATCTGTTAGGCTTGGGAGCACCCTGATCCATGTTGTGATCAA 780
QY 1161 GTATGCTTACCTAGGAGGATGACAGAGCCATGATGATGATGATGATGATGATGAT 1220
DB 781 GTATGCTTACCTAGGAGGATGACAGAGCCATGATGATGATGATGATGATGATGAT 840
QY 1221 CA 1222
DB 841 CA 842

RESULT 14
CB973407
LOCUS
DEFINITION CB973407 867 bp mRNA linear EST 01-MAY-2003
vitisfera cDNA clone CAB30002_Ic_Fc_D10 5', mRNA sequence.
CB973407
ACCESSION
VERSION CB973407.1 GI:30296613
KEYWORDS
SOURCE
ORGANISM
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
1 (bases 1 to 867)
Jones, K. and Cook, D.
Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
berries at various developmental stages
Unpublished (2003)
Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcoc@ucdavis.edu
Seq primer: ACGGTACCGACATATGCC.
Location/Qualifiers

FEATURES
source
1..867
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultiivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CAB30002_Ic_Fc_D10"
/sex="Hermaphrodite"
/dev_stage="Berry stage I"
/lab_host="DH5alpha"
/clone_lib="Cabernet Sauvignon Berry Stage I - CAB3"

/note="Organ: Berry; Vector: pDNR; Site 1: sf1; Site 2: sf1; CAB3 is a cDNA library of Vitis vinifera 'Cabernet Sauvignon' clone 8 berries. Samples were collected after berry set from field-grown vines during stage 1 of berry growth, 17 days after full bloom. The average berry size was 6 millimeters. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCATGTGATACGACAGTGGCCATTCAGCCGG-3' and 5'-ATTCTAGAGCCGAGCGGCGGACGATG-3' (30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

```

Query Match      40.7%; Score 623.4; DB 6; Length 867;
Best Local Similarity 82.4%; Pred. No. 4e-177;
Matches 714; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 55 GAGATTCACTCCGTGACAACTACGAAACCAACCGAATTGTCACGAAACGAAATGGC 114
DB 1 GAGATTCACTCCGTGACAACTACGAAACCAACCGAATTGTCACGAAACGAAATGGC 60
QY 115 ACCTATCACTGATGTCACAAACCAATCCGTCACTACCAATTAAACCAACCCAT 174
DB 61 ACCTATCACTGATGTCACAAACCAATCCGTCACTACCAATTAAACCAACCCAT 120
QY 175 GTTCCAAATTTGGGGGTGATGCTTGTGGTGGGGTGAACAACCGCTCTACCTCACC 234
DB 121 GTTCCCAATTTGGGGGTGATGCTTGTGGTGGGGTGAACAACCGCTCTACCTCACC 180
QY 235 GGTGGTGTATTGCTTAACAGAGAGCATTTTCATGGGCTTACAAAGACAAAGATTCAACA 294
DB 181 GGTGGTGTATTGCTTAACAGAGAGCATTTTCATGGGCTTACAAAGACAAAGATTCAACA 240
QY 295 GCCAATTACTTTGGCTCCCTCAACCCAGCCTCAGCTATTCAGAGTTGATCTTCCAGGGA 354
DB 241 GCCAATTACTTTGGCTCCCTCAACCCAGCCTCAGCTATTCAGAGTTGATCTTCCAGGGA 300
QY 355 GAGAAATCTATGCCCATTCACAAAGTCTGCTTCCAAATGTTAATCTGACGACATTTGTG 414
DB 301 GAGGAGATTATGAGCTTCATTAAGAGCTCTCTCCCTAATGGTAAACCAAGATGACATTTGT 360
QY 415 TTTGGGGGATTTGGGATATGCAACATGAACCTGGCTGATAGCCATGGCCGAGGCAAGGTG 474
DB 361 TTTGGGGGATTTGGGATATGCAACATGAACCTGGCTGATAGCCATGGCCGAGGCTAAGGTC 420
QY 475 TTGACATCATTTTGGAGAGAGTTGAGGCTTACATGGAATCCATGGTCCACTCCCC 534
DB 421 CTGGACATTTGATCTGACAGAGAGTTGAGGCTTACATGGAATCCATGGTCCACTCCCC 480
QY 535 GGAATCTACGACCCGAGATTTCATTTGCTGCCAACAGAGAGAGCTGCCAACACGTGAT 594
DB 481 GGAATCTACGACCCGAGATTTCATTTGCTGCCAACAGAGAGAGCTGCCAACACGTGAT 540
QY 595 AAGGGCAACAAGACAGAGAGTTGACGAATCATCAAGACATCAAGGCTTTAAGAA 654
DB 541 AAGGGCAACAAGACAGAGAGTTGACGAATCATCAAGACATCAAGGCTTTAAGAA 600
QY 655 GCACCAAGGTGAGACAGGTGTTGCTGTGGAATGCAACAGAGAGATATGCAAT 714
DB 601 GCACCAAGGTGAGATGAGGTGTTGCTGTGGAATGCAACAGAGAGATATGCAAT 660
QY 715 TTGTTGTTAGGCTTAATGACACATGAGAAATCTTGGCTGCTGTGACAGAAATGAG 774
DB 661 TTGTTGTTAGGCTTAATGACACATGAGAAATCTTGGCTGCTGTGACAGAAATGAG 720
QY 775 GGTGAGATTTCTCTTCCACCTTTGATGCAATTTGCTGTGATGGAATATGTTCTTTTC 834
DB 721 CCGTGAATCTCTCTTCCACCTTTGATGCAATTTGCTGTGATGGAATATGTTCTTTTC 780
QY 835 ATTATGGAAGCCCTCAGAACTTTTGAACAGGAGGCTATGATCTTGCATGCGAGG 894

```

```

DB 781 ATCATGAGAGCCACAGAACACTTTTGTCCAGGCGTATGATTGGCTATTAAAGG 840
QY 895 AACACTTTGATTGTGGAGATGACTT 920
DB 841 AACAGTTTGAATTGTGGGATGATATT 866

```

RESULT 15

CP212628

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CP212628 891 bp mRNA linear EST 01-AUG-2003
 CGF1000655 A10 Vitis vinifera cv. cabernet sauvignon Stem - CAST
 Vitis vinifera cDNA clone CAST0005_11IF_A10 5', mRNA sequence.
 CP212628
 CF212628.1 GI:33407001
 EST.
 Vitis vinifera
 Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 891)
 Goes da Silva, F., Iandolo, A., Jim, H., Baek, J., Leslie, A., Xu, J.,
 Jones, R., and Cook, D.
 Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
 berries at various developmental stages
 Unpublished (2003)
 Contact: Douglas Cook, PhD
 CAES Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcoc@ucdavis.edu
 Seq primer: ACGTACCGACATATGCC.
 Location/Qualifiers
 1..891
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CAST0005_11IF_A10"
 /sex="Hermaphrodite"
 /dev_stage="Pre-bloom (10-11 days before bloom)"
 /lab_host="DH5alpha"
 /clone_11b="Vitis vinifera cv. cabernet sauvignon Stem -
 CAST"

/note="Organ: Stem; Vector: pDNR; Site 1: sf1; Site 2: sf1; CAST is a cDNA library of Vitis vinifera cv. 'Cabernet Sauvignon' stems. Samples were collected May 13, 2002 from pre-bloom plants (10-11 days before bloom), pre-veraison. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCATGTGATACGACAGTGGCCATTCAGCCGG-3' and 5'-ATTCTAGAGCCGAGCGGCGGACGATG-3' (30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

```

Query Match      40.5%; Score 620.4; DB 7; Length 891;
Best Local Similarity 83.6%; Pred. No. 3.3e-176;
Matches 702; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 637 ATCAAGGCGTTTAAAGAGCCCAAGATGACAAAGGTGTTGCTGTGACTGCCAAC 696
DB 51 ATTAGGAGCTTTAAGAGAGCAACAGGTGATTAAGTGTGCTGTGACTGCCAAC 110
QY 697 ACAGAGAGTATAGCAATTTGGTGTAGAGCTTAAATGACACCATGAGAAATCTTGGCT 756
DB 111 ACAGAGCATACAGTAATTTGATTTGGAGCTGAATGACACCATGAGAAATCTTGGCT 170
QY 757 GCTGTGACAGAAATAGAGCTGAGATTTCTCTTCCACTTGTATGCAATGGCCGTGTG 816

```

```
Db      171 TCTCTGGATAGGAATAGACCTGAGATCTCTCCACCTGTATGCTCTGCTGTGTC 230
Qy      817 ATGGAATATGTTCTCTTATTATAGAAAGCCCTCAGAACCTTTGTACCAAGGCTGAT 876
Db      231 ATGGAATATGTTCTCTTATTATAGAAAGCCCTCAGAACCTTTGTTCAGGCTGATC 290
Qy      877 GATCTTGCATCGCAGGAACACTTGTATGTGTGAGATGACTTCAGAGTGTGCAGAC 936
Db      291 GATTGGCTATTAAGAGAGACGTTTGTATGGTGGGATGATTTTAAGATGTGTCAAAC 350
Qy      937 AAAATGAATCTGTGTGTGATTTTCTGTGGGGCTGTGTATCAAGCCAACATCTATA 996
Db      351 AAGATGAATCTGTATGTGTGATTTCTGTGTGGGCAAGCATCAAGCCAACATCAAT 410
Qy      997 GTTAGTTCAACCATCTGGGAAACAATGTGTATGATCTCTGGCTCCACAACCTTC 1056
Db      411 GTAGCTATTAACCATCTGGGAAACAATGTGTATGATCTGTGTCTCTCAAGCCTC 470
Qy      1057 CGCTTCAAGGAATCTCCAGAGCAACGTTGTGAAGATATGTGCACAGCAATGCCATC 1116
Db      471 CGCTCAAGAGATCTCCAGAGCAATGTGTGACGACATGTGTTCCAGCAAGCTATC 530
Qy      1117 CTCTATGAGCTGTGGAACATCTCCGACCATGTGTGTATTAAGTATGTGCTTACGTA 1176
Db      531 CTGTATGAGCTGTGGGAGCACCTGACCATGTGTGTATCAAGTATGTGCATATGTT 590
Qy      1177 GGGATAGCAGAGAGCCATGTATGATGATACCTTCAGAGATATTCATGTGTGAAAGAAC 1236
Db      591 GGAGACAGCAGAGAGCCATGTATGATGATACCTTCAGAGATATTCATGTGTGAAAGAC 650
Qy      1237 ACCATTGTTTTCACACACATGTAGAGATTCCTTTAGCTGCTCTATTAATCTTGAC 1296
Db      651 ACCATTGTGTGCAACAACCTGCGAGAGCTCATGTGTGCTGCCATCATCTTGANT 710
Qy      1297 TTGTCTCTTCTTGCTGAGCTGAGCACTAGATCCAGTTTAAAGCTGAAATGAGGAAAA 1356
Db      711 TTGTCTCTTCTTGCTGAGCTGAGCACCGCATCCAGCTCAAGGCTGAAAGAGGACAA 770
Qy      1357 TTCCACTATTCCACCGAGTGTCTACCATTTCTCAGCTATCTGACCAAGGCTCCTGGTT 1416
Db      771 TTTCATTCTTCCACCGAGTGTCTACCATTTCTCAGTTTACCAAGGCTCCTGGTT 830
Qy      1417 CCACCGGATACACAGTGTGATGATGATTCMAAGCAGCGTGAATGCTGAAAAACATA 1476
Db      831 CCACCGGATACACAGTGTGATGATGATTCMAAGCAGCGTGAATGCTGAAAAACATA 890
```

Search completed: June 8, 2005, 09:51:52
Job time : 5228.81 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 16:00:16 ; Search time 843.546 Seconds

(without alignments)
10758.117 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533

Sequence: 1 agcttcacgcagaattttaa.....tgattctcgagtaacaagcga 1533

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1533	100.0	1533	12	ADQ14500	Adq14500 Mutant so
2	1533	100.0	1533	13	ADS82001	AdS82001 Soybean m
3	1531.4	99.9	1533	12	ADQ14504	Adq14504 Wild type
4	1531.4	99.9	1533	12	ADQ14502	Adq14502 Mutant so
5	1531.4	99.9	1533	13	ADS82005	AdS82005 Soybean m
6	1531.4	99.9	1533	13	ADS82003	AdS82003 Soybean m
7	1464.2	95.5	1533	12	ADQ14498	Adq14498 Mutant so
8	1464.2	95.5	1533	13	ADS81999	AdS81999 Soybean m
9	1464.2	95.5	1533	13	ADS81993	AdS81993 Soybean m
10	1464.2	95.5	1760	12	ADQ14490	Adq14490 Wild type
11	1464.2	95.5	1782	12	AAV62440	AAV62440 Soybean w
12	1462.6	95.4	1533	2	AAV62443	AAV62443 Soybean m
13	1462.6	95.4	1533	2	ADQ14494	Adq14494 Mutant so
14	1462.6	95.4	1533	13	ADS81997	AdS81997 Soybean m
15	1097.8	71.6	1950	2	AAx30402	AAx30402 Nicotiana
16	1000.2	65.2	1781	4	AAc87643	AAc87643 Brassica
17	998.6	65.1	1533	6	ABz13633	ABz13633 Arabidops
18	998.6	65.1	1533	12	ADN73524	ADN73524 Thale cre
19	998.6	65.1	1938	3	AAc50242	AAc50242 Arabidops
20	987.2	64.4	1536	3	AAc34428	AAc34428 Arabidops

21	985.8	64.3	1665	2	AAx09006	AAx09006 Inducible
22	985.6	64.3	1837	3	AAc34806	AAc34806 Arabidops
23	984.2	64.2	1959	4	AAc85922	AAc85922 MIP synch
24	979.4	63.9	1931	2	AAx24407	AAx24407 Maize myo
25	941.6	61.4	1759	6	ABq72653	ABq72653 Human MDD
26	821.4	53.6	1536	12	ADP43918	ADP43918 P. coarct
27	782	51.0	1719	3	AAc51095	AAc51095 Arabidops
28	547.2	35.7	1605	8	ABT19345	ABT19345 Aspergill
29	547.2	35.7	1605	8	ABT21165	ABT21165 Aspergill
30	527.2	34.4	2280	4	ABL17361	ABL17361 Drosophill
31	499	32.6	1772	4	AAH14171	AAH14171 Human CDN
32	499	32.6	1818	4	AAH15251	AAH15251 Human CDN
33	499	32.6	1825	6	ABT07176	ABT07176 Human ova
34	499	32.6	1852	4	AAH47741	AAH47741 Human hMI
35	499	32.6	1862	4	AAI59232	AAI59232 Human pol
36	499	32.6	2380	13	ADQ86789	ADQ86789 Human tum
37	499	32.6	2380	13	ADQ85720	ADQ85720 Tumour-as
38	499	32.6	2380	13	ACN40520	ACN40520 Human CDN
39	486.6	31.7	1991	12	ADG73735	ADG73735 Aspergill
40	478.2	31.2	1833	4	AAH47740	AAH47740 Murine mM
41	464.4	30.3	618	13	ACN62990	ACN62990 Cotton ca
42	453.6	29.6	1835	12	ADQ25429	ADQ25429 Human sof
43	451.2	29.4	1702	10	ADB62765	ADB62765 Human CDN
44	450.8	29.4	2338	4	AAI61018	AAI61018 Human pol
45	434.2	28.3	1704	8	ABT20567	ABT20567 Aspergill

ALIGNMENTS

RESULT 1	
ID ADQ14500	standard; cDNA; 1533 BP.
XX	
AC ADQ14500;	
XX	
DT 23-SEP-2004	(first entry)
XX	
DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #3.	
XX	
KW soybean; myo-inositol 1-phosphate synthase; gene; ss;	
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;	
KW inorganic phosphate; mutant.	
XX	
OS Glycine max.	
OS Synthetic.	
XX	
FT Key	Location/Qualifiers
FT CDS	1..1533
FT	/*tag= a
FT	/product= "Mutant soybean myo-inositol 1-phosphate synthase #3"
XX	
PN US2004128713-A1.	
XX	
PD 01-JUL-2004.	
XX	
PF 21-NOV-2003; 2003US-00718952.	
XX	
PR 08-APR-1997; 97US-00835751.	
PR 07-APR-1998; 98MO-US006822.	
PR 26-APR-1999; 99US-0029315.	
PR 11-MAR-2002; 2002US-00025003.	
XX	
PA (HITZ/) HITZ W D.	
PA (SEBA/) SEBASTIAN S A.	
PA (GRAC/) GRACE D J.	
PA (STRE/) STREIT L G.	
XX	
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;	
XX	
DR WPI: 2004-533135/51.	
DR P-PSDB; ADQ14501.	

XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

PS Claim 10; SEQ ID NO 11; 48bp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

Query Match 100.0%; Score 1533; DB 12; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGTCATGAGAAATTTTAAAGTAGAGTCTTAATGTGAATACCGAGACTGAGATT 60
DB 1 AAGTCATGAGAAATTTTAAAGTAGAGTCTTAATGTGAATACCGAGACTGAGATT 60
QY 61 CAGTCCGCTGATACACTACGAAACACCGAATTTGTTCAAGAAACAGAAATGCACTAT 120
DB 61 CAGTCCGCTGATACACTACGAAACACCGAATTTGTTCAAGAAACAGAAATGCACTAT 120
QY 121 CAGTGGATTTGCAAAACCCAAATCCGTCACTACCAATTTAAACCAACCAATGTTCCA 180
DB 121 CAGTGGATTTGCAAAACCCAAATCCGTCACTACCAATTTAAACCAACCAATGTTCCA 180
QY 121 CAGTGGATTTGCAAAACCCAAATCCGTCACTACCAATTTAAACCAACCAATGTTCCA 180
DB 121 CAGTGGATTTGCAAAACCCAAATCCGTCACTACCAATTTAAACCAACCAATGTTCCA 180
QY 181 AAATTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAGGCTTACCCCTGACCGGTGAT 240
DB 181 AAATTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAGGCTTACCCCTGACCGGTGAT 240
QY 181 AAATTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAGGCTTACCCCTGACCGGTGAT 240
DB 181 AAATTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAGGCTTACCCCTGACCGGTGAT 240
QY 241 GTTATTTGCTAACAGAGAGACATTTTCATGGGCTAACAAAGACAGATTCACCAAGCCAT 300
DB 241 GTTATTTGCTAACAGAGAGACATTTTCATGGGCTAACAAAGACAGATTCACCAAGCCAT 300
QY 301 TACTTTGGCTCCCTCCACCAAGGCTCAGCTATTCAGATTGATCTTCCAGGAGAGGAA 360
DB 301 TACTTTGGCTCCCTCCACCAAGGCTCAGCTATTCAGATTGATCTTCCAGGAGAGGAA 360
QY 301 TACTTTGGCTCCCTCCACCAAGGCTCAGCTATTCAGATTGATCTTCCAGGAGAGGAA 360
DB 301 TACTTTGGCTCCCTCCACCAAGGCTCAGCTATTCAGATTGATCTTCCAGGAGAGGAA 360
QY 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAATGTTTAACTCTGACACATTTGTTGGG 420
DB 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAATGTTTAACTCTGACACATTTGTTGGG 420
QY 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAATGTTTAACTCTGACACATTTGTTGGG 420
DB 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAATGTTTAACTCTGACACATTTGTTGGG 420
QY 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCAAGGCGCAAGGAGGTTTGAAC 480
DB 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCAAGGCGCAAGGAGGTTTGAAC 480
QY 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCAAGGCGCAAGGAGGTTTGAAC 480
DB 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCAAGGCGCAAGGAGGTTTGAAC 480
QY 481 ATCGATTTTCAGAGAGAGTGAAGCTTACATGAATTCATGGTTTCACTCCCGGAATC 540
DB 481 ATCGATTTTCAGAGAGAGTGAAGCTTACATGAATTCATGGTTTCACTCCCGGAATC 540
QY 481 ATCGATTTTCAGAGAGAGTGAAGCTTACATGAATTCATGGTTTCACTCCCGGAATC 540
DB 481 ATCGATTTTCAGAGAGAGTGAAGCTTACATGAATTCATGGTTTCACTCCCGGAATC 540
QY 541 TACGACCCGAGATTTCACTGCTGCAACCAAGAGAGCGTCCCAACAGTGAATTAAGGCG 600
DB 541 TACGACCCGAGATTTCACTGCTGCAACCAAGAGAGCGTCCCAACAGTGAATTAAGGCG 600

QY 601 ACAAAGCAGAGCAAGTTACAGCAATATCATGAAGATCAAGCGCTTTAAGAAACCAAC 660
DB 601 ACAAAGCAGAGCAAGTTACAGCAATATCATGAAGATCAAGCGCTTTAAGAAACCAAC 660
QY 661 AAAGTGACAAAGGTGTTGCTCTGTGACCTGCCAACACAGAGAGTATAGAAATTTGGTT 720
DB 661 AAAGTGACAAAGGTGTTGCTCTGTGACCTGCCAACACAGAGAGTATAGAAATTTGGTT 720
QY 661 AAAGTGACAAAGGTGTTGCTCTGTGACCTGCCAACACAGAGAGTATAGAAATTTGGTT 720
DB 661 AAAGTGACAAAGGTGTTGCTCTGTGACCTGCCAACACAGAGAGTATAGAAATTTGGTT 720
QY 721 GTAGGCTTTATGACACCATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTTATGACACCATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTCCACCTTGTATGATGCAATGCTGTGTGATGAGAAATGTTCTTTCAATAT 840
DB 781 ATTTCTCTCCACCTTGTATGATGCAATGCTGTGTGATGAGAAATGTTCTTTCAATAT 840
QY 841 GGAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTGGCATGCGAGAGAACCT 900
DB 841 GGAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTGGCATGCGAGAGAACCT 900
QY 901 TTGATTTGGTGGAGATGACTTCAAGAGTGTGACAGCAACAAATGAAATCTGTGTGGTAT 960
DB 901 TTGATTTGGTGGAGATGACTTCAAGAGTGTGACAGCAACAAATGAAATCTGTGTGGTAT 960
QY 961 TTTCTTGTGGGGGCTGTATACAGCAACATCTATAGTTAGTTACACCATCTGGGAAAC 1020
DB 961 TTTCTTGTGGGGGCTGTATACAGCAACATCTATAGTTAGTTACACCATCTGGGAAAC 1020
QY 1021 AATGATGATGATATCTCTGGCTCCACAAACCTTCCGCTCCAGGAATCTCCAGAGAC 1080
DB 1021 AATGATGATGATATCTCTGGCTCCACAAACCTTCCGCTCCAGGAATCTCCAGAGAC 1080
QY 1081 AAGCTTTGAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 AAGCTTTGAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 GACCAATGTTGTTTATTAAGTATGCTCTTACGTAAGGAGATGACAGAGACCATGAT 1200
DB 1141 GACCAATGTTGTTTATTAAGTATGCTCTTACGTAAGGAGATGACAGAGACCATGAT 1200
QY 1201 GAGTACATTTCAAGATATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 GAGTACATTTCAAGATATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 GAGATTCCTTTTATGCTGCTCTTATCTTTGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 GAGATTCCTTTTATGCTGCTCTTATCTTTGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 ACTAGAAATTCAGTTTAAAGCTGAAATGAGGAGAAATTCACCTCATTCACCCAGTTGCT 1380
DB 1321 ACTAGAAATTCAGTTTAAAGCTGAAATGAGGAGAAATTCACCTCATTCACCCAGTTGCT 1380
QY 1381 ACCATTTCTCAGTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACACAGTGTGAT 1440
DB 1381 ACCATTTCTCAGTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACACAGTGTGAT 1440
QY 1441 GCATTTGCAAGCAGCTGCAATGCTGGAATAACATATGAGGGCTTGTGTTGATGAGCC 1500
DB 1441 GCATTTGCAAGCAGCTGCAATGCTGGAATAACATATGAGGGCTTGTGTTGATGAGCC 1500
QY 1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533
DB 1501 CCAGAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533

RESULT 2
ADSS82001
ID ADSS82001 standard; cDNA; 1533 BP.
XX
AC ADSS82001;
XX
DT 18-NOV-2004 (first entry)
XX

DE Soybean myo-inositol 1-phosphate synthase cDNA mutant #2.
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
XX
OS Glycine max; line 29010CP01.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
FT mutation replace(260,G)
FT /tag= b
PN US2003074695-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
XX
PR 07-APR-1998; 98WO-US006822.
XX
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
PI Hitz WD, Sebastian SA;
XX
XX WPI; 2004-639957/62.
XX P-PSDB; ADS82002.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 8; SEQ ID NO 11; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC phosphate synthase having decreasing capacity for the synthesis for myo-
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC complement, subfragment or the complement of the subfragment, operably
CC linked to suitable regulatory sequences, where expression of the chimeric
CC gene results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC from the plant, making a soybean plant with the heritable phenotype
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
CC an elite soybean plant and selecting a progeny plant of the cross of
CC crossing step that has a heritable phenotype as mentioned above), seeds
CC of soybean plant made by the above method, a soy protein product derived
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC heritable phenotype as mentioned above), and making or producing a
CC soybean protein product derived from seeds of a soybean plant with a
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a mutant myo-
CC inositol 1-phosphate synthase.
XX
XX
SQ Sequence 1533 BP; 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

Query Match 100.0%; Score 1533; DB 13; Length 1533;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAATACACCGAGCTGAGATT 60
|||
DB 1 ATGTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAATACACCGAGCTGAGATT 60
QY 61 CAGTCGNGTACAACTAGAAACACCGAATCTTGTTCACGAAACAGAAATGGACCTAT 120
|||
DB 61 CAGTCGNGTACAACTAGAAACACCGAATCTTGTTCACGAAACAGAAATGGACCTAT 120
QY 121 CAGTGGATTGTCAAAACCCAAATCCGTCACTACCAATTTTAAACCAACACCCATGTTCCA 180
|||
DB 121 CAGTGGATTGTCAAAACCCAAATCCGTCACTACCAATTTTAAACCAACACCCATGTTCCA 180
QY 181 AAATTGGGGGTGATGCTTGTGGGTTGGGGTGGAAACACCGCTCTACCTCACCGGTGT 240
|||
DB 181 AAATTGGGGGTGATGCTTGTGGGTTGGGGTGGAAACACCGCTCTACCTCACCGGTGT 240
QY 241 GTTATTTCTTAACAGAGAGACATTTCAATGGGCTTACAAAGGAGCAAGATTTCAACAGCCAT 300
|||
DB 241 GTTATTTCTTAACAGAGAGACATTTCAATGGGCTTACAAAGGAGCAAGATTTCAACAGCCAT 300
QY 301 TACTTTGGCTCCCTCAGACCCAAAGCTCAGCTATTTCAGATTGATTCCTTCCAGGAGAGAA 360
|||
DB 301 TACTTTGGCTCCCTCAGACCCAAAGCTCAGCTATTTCAGATTGATTCCTTCCAGGAGAGAA 360
QY 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCATATGTTTATCTTGACGACATTTGTGGG 420
|||
DB 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCATATGTTTATCTTGACGACATTTGTGGG 420
QY 421 GGATGGGATATAGCAAACTGAACCTGGCTGATGTGCATGGCCAGGAGCAAAAGGTTTTC 480
|||
DB 421 GGATGGGATATAGCAAACTGAACCTGGCTGATGTGCATGGCCAGGAGCAAAAGGTTTTC 480
QY 481 ATCGATTTGCAAGAACAGTTGAGGCTTACATGATCATGATCGATGTTTCCATCCCGGAATC 540
|||
DB 481 ATCGATTTGCAAGAACAGTTGAGGCTTACATGATCATGATCGATGTTTCCATCCCGGAATC 540
QY 541 TACGACCCGGATTTTCAATGCTGCTCCAAACAGAGGAGCGGCCAAACAGTGATTAAGGCG 600
|||
DB 541 TACGACCCGGATTTTCAATGCTGCTCCAAACAGAGGAGCGGCCAAACAGTGATTAAGGCG 600
QY 601 ACAAAGCAAGAGCAAGTTCAGCAATCATCAAAAGCATCAAGAGCGTTTAAAGAAACCCAC 660
|||
DB 601 ACAAAGCAAGAGCAAGTTCAGCAATCATCAAAAGCATCAAGAGCGTTTAAAGAAACCCAC 660
QY 661 AAAGTGACAAAGGTGGTGTCTGTGACCTGCAACACAGAGAGGTATAGCAATTTGGTT 720
|||
DB 661 AAAGTGACAAAGGTGGTGTCTGTGACCTGCAACACAGAGAGGTATAGCAATTTGGTT 720
QY 721 GTAGGCTTTAATGACACCATGAGAAATCTTTGCTGTGTGACAGAAATGAGGCTGAG 780
|||
DB 721 GTAGGCTTTAATGACACCATGAGAAATCTTTGCTGTGTGACAGAAATGAGGCTGAG 780
QY 781 ATTTTCTCTTCCACTTGTATGACCATTTGCTGTGTGATGAGAAATGTTCTTCAATTAAT 840
|||
DB 781 ATTTTCTCTTCCACTTGTATGACCATTTGCTGTGTGATGAGAAATGTTCTTCAATTAAT 840
QY 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCATGCGAGAACACT 900
|||
DB 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCATGCGAGAACACT 900
QY 901 TTGATTTGAGATGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTTGAT 960
|||
DB 901 TTGATTTGAGATGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTTGAT 960
QY 961 TTTCTTGTGGGGGCTGTGTATCAAGCAACATCTAATAGTTAATCAACCATCTGGGAAAC 1020
|||
DB 961 TTTCTTGTGGGGGCTGTGTATCAAGCAACATCTAATAGTTAATCAACCATCTGGGAAAC 1020
QY 1021 AATGATGATATGAAATCTCTCGGCTCCAAACCTTCCGCTCCAGAGAAATCTCCAAAGAC 1080
|||
DB 1021 AATGATGATATGAAATCTCTCGGCTCCAAACCTTCCGCTCCAGAGAAATCTCCAAAGAC 1080

QY 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGAACATCCC 1140
DB 1081 AACGTTGTGACGATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGAACATCCC 1140
QY 1141 GACCATGTTGTTTATTTAAGTATGCTTACGTAAGGAGATACGAAAGACCATGAT 1200
DB 1141 GACCATGTTGTTTATTTAAGTATGCTTACGTAAGGAGATACGAAAGACCATGAT 1200
QY 1201 GAGTACACTTCAGATATTCATGGGTGGAAGAACACCATTTGTCACAACATGT 1260
DB 1201 GAGTACACTTCAGATATTCATGGGTGGAAGAACACCATTTGTCACAACATGT 1260
QY 1261 GAGGATTCCTTTTACCTCTCTATTTATTTGACATTTGCTCTTTGCTGAGCTGAC 1320
DB 1261 GAGGATTCCTTTTACCTCTCTATTTATTTGACATTTGCTCTTTGCTGAGCTGAC 1320
QY 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACATTCATCCACCATTTGCT 1380
DB 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACATTCATCCACCATTTGCT 1380
QY 1381 ACCATTCAGCTATCTGACCAAGGCTCTGTTCCACCGGGTACACGAGTGTGAT 1440
DB 1381 ACCATTCAGCTATCTGACCAAGGCTCTCTGTTCCACCGGGTACACGAGTGTGAT 1440
QY 1441 GCATTGTCAAAGCAGCTGCAATGTCTGAAACATTAATGAGGCTTTGTTGATTTGGCC 1500
DB 1441 GCATTGTCAAAGCAGCTGCAATGTCTGAAACATTAATGAGGCTTTGTTGATTTGGCC 1500
QY 1501 CCAGAGATTAACATGATTTCTCGATACAACTGA 1533
DB 1501 CCAGAGATTAACATGATTTCTCGATACAACTGA 1533

RESULT 3
ADQ14504
ID ADQ14504 standard; cDNA; 1533 BP.
AC ADQ14504;
XX

23-SEP-2004 (first entry)

Wild type soybean myo-inositol 1-phosphate synthase cDNA #2.

Soybean; myo-inositol 1-phosphate synthase; gene; ss;
myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
inorganic phosphate.

Glycine max.

Key Location/Qualifiers
CDS 1..1533
FT /tag= a
FT /product= "Wild type soybean myo-inositol 1-phosphate
FT synthase #2"
XX

US2004128713-A1.

01-JUL-2004.

21-NOV-2003; 2003US-00718952.

08-APR-1997; 97US-00835751.

07-APR-1998; 98MC-US006822.

26-APR-1999; 99US-00299315.

11-MAR-2002; 2002US-00025003.

(HITZ/) HITZ W D.
(SEBA/) SEBASTIAN S A.
(GRAC/) GRACE D J.
(STRE/) STREIT L G.

Hitz WD, Sebastian SA, Grace DJ, Streit LG;

XX WPI; 2004-533135/51.
DR P-PsDB; ADQ14505.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
PS Claim 4; SEQ ID NO 15; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 99.9%; Score 1531.4; DB 12; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTTCATCGAGATTTTAAAGTAGAGAGTCTTAATGTAAGTACCCGAGATCGAATT 60
DB 1 ATGTTTCATCGAGATTTTAAAGTAGAGAGTCTTAATGTAAGTACCCGAGATCGAATT 60
QY 61 CAGTCGGTGAACACTAGAGAAACCCGAACTGTTTCAGAGAACGAAATGACCCTAT 120
DB 61 CAGTCGGTGAACACTAGAGAAACCCGAACTGTTTCAGAGAACGAAATGACCCTAT 120
QY 121 CAGTGGATTTGCAAAACCCAAATCCGTCACCTAACCAATTTAAACCAACCATTTTCCA 180
DB 121 CAGTGGATTTGCAAAACCCAAATCCGTCACCTAACCAATTTAAACCAACCATTTTCCA 180
QY 181 AAATTGGGGGTGATGCTTGTGGGTGGGTAACCAACGCTCTACCTTACCGGTGT 240
DB 181 AAATTGGGGGTGATGCTTGTGGGTGGGTAACCAACGCTCTACCTTACCGGTGT 240
QY 241 GTTATTTGTAACAGAGAGACATTTTCATGGGCTCAAGAGCAAGATTCACCAAGCAAT 300
DB 241 GTTATTTGTAACAGAGAGACATTTTCATGGGCTCAAGAGCAAGATTTTCACCAAGCAAT 300
QY 301 TACTTTGGCTCCCTACCCCAAGCTCAGCTATTGAGTTGGATCTCTTCAAGGAGAGAA 360
DB 301 TACTTTGGCTCCCTACCCCAAGCTCAGCTATTGAGTTGGATCTCTTCAAGGAGAGAA 360
QY 361 ATCTATGCCCATTCGAAGAGTCTGCTCAATGTTAACTCTGACGACATTTGTTGGG 420
DB 361 ATCTATGCCCATTCGAAGAGTCTGCTCAATGTTAACTCTGACGACATTTGTTGGG 420
QY 421 GAGTGGGATTTACGAACATGAACCTGGCTGATGCAAGGCAAGGATGTTTAC 480
DB 421 GAGTGGGATTTACGAACATGAACCTGGCTGATGCAAGGCAAGGATGTTTAC 480
QY 481 ATCGATTTGCAAGACAGTTGAGGCTTTACATGGAATCCATGTTTCCACTCCCGGAATC 540
DB 481 ATCGATTTGCAAGACAGTTGAGGCTTTACATGGAATCCATGTTTCCACTCCCGGAATC 540
QY 541 TACGACCCGGAATTTGATTTGCTGCCAACCAAGAGAGCGTGCCCAACAGTATTAAGGCG 600

Db	541	TACGACCCGGATTTCATTGCTGCGACCAAGAGAGCGCTGCCAACACGTGATTAAAGGC	600
Qy	601	ACAAAGCAAGCAAGCAAGTTCAAGAAATCATCAAGACATCAAGGCGTTTAAAGAACCAAC	660
Db	601	ACAAAGCAAGCAAGCAAGTTCAAGAAATCATCAAGACATCAAGGCGCTTTAAAGAACCAAC	660
Qy	661	AAAGTGACAAAGTGAGTGTCTGTGGACTGGCAACACAGAGAGGATTAGCAATTTGGTT	720
Db	661	AAAGTGACAAAGTGAGTGTCTGTGGACTGGCAACACAGAGAGGATTAGCAATTTGGTT	720
Qy	721	GTAGGCGCTTAAATGACACCATGAGAAATCTTTGCGTGTGGAAGAATGAGGCTAG	780
Db	721	GTAGGCGCTTAAATGACACCATGAGAAATCTTTGCGTGTGGAAGAATGAGGCTAG	780
Qy	781	ATTTCCTCTTCAACCTTGATGCGCATTCGCTGTGATGAGAAAAGTTCCTTCAATAT	840
Db	781	ATTTCCTCTTCAACCTTGATGCGCATTCGCTGTGATGAGAAAAGTTCCTTCAATAT	840
Qy	841	GGAAGCCCTCAGAACACTTTTGTACCAAGGCGATTTGATCTTTGCATCGCAGAGAACCT	900
Db	841	GGAAGCCCTCAGAACACTTTTGTACCAAGGCGATTTGATCTTTGCATCGCAGAGAACCT	900
Qy	901	TTGATTTGATGAGATGATCTTCAAGAGTGCTCAGACCAACAAATGAAATCTGTGTTGGAT	960
Db	901	TTGATTTGATGAGATGATCTTCAAGAGTGCTCAGACCAACAAATGAAATCTGTGTTGGAT	960
Qy	961	TTTCTTGTGGGGGCGGTATCAGCAACATCTATAGTTAGTTACCAACATCTGGGAAAC	1020
Db	961	TTTCTTGTGGGGGCGGTATCAGCAACATCTATAGTTAGTTACCAACATCTGGGAAAC	1020
Qy	1021	AATGATGATGAAATCTCTCGGCTCCACAAACCTTCGCGCTCAAGAAATCTCCAAAGC	1080
Db	1021	AATGATGATGAAATCTCTCGGCTCCACAAACCTTCGCGCTCAAGAAATCTCCAAAGC	1080
Qy	1081	AAAGTGTGGAAGATAGTGTCAACAGCAATGCAATCCCTTATGAGCGCTGTGAAACATGCC	1140
Db	1081	AAAGTGTGGAAGATAGTGTCAACAGCAATGCAATCCCTTATGAGCGCTGTGAAACATGCC	1140
Qy	1141	GACCATGTTGTTGTTATTAAAGTATGTGCTTACGTAGGGGATAGCAGAGAGCCATGGAT	1200
Db	1141	GACCATGTTGTTGTTATTAAAGTATGTGCTTACGTAGGGGATAGCAGAGAGCCATGGAT	1200
Qy	1201	GAGTACACTTCAAGATATTTCATGGGTGGAAGAACACCATTTGTTTTCACAACACATGT	1260
Db	1201	GAGTACACTTCAAGATATTTCATGGGTGGAAGAACACCATTTGTTTTCACAACACATGT	1260
Qy	1261	GAGGATTCCTTTTAAAGCTCTTATTAATCTTGAACCTTGTCCTTCTTGCTGAGCTAGC	1320
Db	1261	GAGGATTCCTTTTAAAGCTCTTATTAATCTTGAACCTTGTCCTTCTTGCTGAGCTAGC	1320
Qy	1321	ACTAGAAATCCAGTTTAAAGCTGAAATAGGGGAAATTCACATCCTACCCAGTTGCT	1380
Db	1321	ACTAGAAATCCAGTTTAAAGCTGAAATAGGGGAAATTCACATCCTACCCAGTTGCT	1380
Qy	1381	ACCAATTCACACTATCTBACCAAGGCTCTCTGGTTCCACGGGGTACCAAGTGTGAT	1440
Db	1381	ACCAATTCACACTATCTBACCAAGGCTCTCTGGTTCCACGGGGTACCAAGTGTGAT	1440
Qy	1441	GCATTGTCAAAAGCAGCGTGAATGCTGAAAAACAATATGAGGGCTTGTGTGATTTGCC	1500
Db	1441	GCATTGTCAAAAGCAGCGTGAATGCTGAAAAACAATATGAGGGCTTGTGTGATTTGCC	1500
Qy	1501	CCAGAGAAATACATGATTTCTGAGATCAAGTGA	1533
Db	1501	CCAGAGAAATACATGATTTCTGAGATCAAGTGA	1533

RESULT 4	
ADQ14502	
ID	ADQ14502 standard; cDNA, 1533 BP.
XX	
AC	ADQ14502;

XX	
DT	23-SEP-2004 (first entry)
DE	Mutant soybean myo-inositol 1-phosphate synthase cDNA #4.
XX	
KW	Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KV	myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KX	inorganic phosphate; mutant.
OS	Glycine max.
OS	Synthetic.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1533
FT	/+tag= a
FT	/product= "Mutant soybean myo-inositol 1-phosphate
FT	synthase #4"
XX	
PN	US2004128713-A1.
XX	
PD	01-JUL-2004.
XX	
PE	21-NOV-2003; 2003US-00718952.
PF	
PR	08-APR-1997; 97US-00835751.
PR	07-APR-1998; 98WO-US006822.
PR	26-APR-1999; 99US-00299315.
PR	11-MAR-2002; 2002US-00025003.
XX	
PA	(HITZ/) HITZ W D.
PA	(SEBA/) SEBASTIAN S A.
PA	(GRAC/) GRACE D J.
PA	(STRE/) STREIT L G.
XX	
PI	Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX	
DR	WPI; 2004-533135/51.
DR	P-PSDB; ADQ14503.
PT	New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT	useful for producing plants with decreased raffinose, stachyose, and
PT	phytic acid and increased sucrose, leading to valuable and useful soybean
PT	products.
XX	
PS	Example 8; SEQ ID NO 13; 48pp; English.
XX	
CC	The invention relates to a nucleic acid fragment encoding a soybean myo-
CC	inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC	synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC	phosphate. The invention also relates to a chimeric gene operably linked
CC	to suitable regulatory sequences, where expression of the chimeric gene
CC	results in a decrease in expression of an endogenous or native gene
CC	encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC	comprising the chimeric gene, a method of making the soybean plant, a
CC	seed of the soybean plant, a soy protein product derived from the
CC	processing of soybean seeds, a method of making or producing a soy
CC	protein product and a method of using a soybean plant homozygous for at
CC	least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC	decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC	nucleic acid fragment and methods are useful for producing plants with
CC	decreased raffinose, stachyose and phytic acid content and increased
CC	sucrose and inorganic phosphate content, leading to valuable and useful
CC	soybean products. This sequence represents cDNA encoding a mutant soybean
CC	myo-inositol 1-phosphate synthase polypeptide of the invention.
XX	
XQ	Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match	99.9%	Score 1531.4	DB 12	Length 1533
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1532	Conservative	0	Mismatches 1	Indels 0
				Gaps 0

0x 1 ATGTTTCATCGAAGATTTTAAGTGTGAGAGTCTTAATGTGAAGTACCCGAGACGTGAATT 60
 |||||

```
Dh 1 ATGTCATCGAAGATTTTAAAGGTAGAGATCCTAAATGTGAAGTACACCGAGCTGAGAT 60
Qy 61 CAGTCGGGTATCAACTAGGAACACCGAATCTGTATCAAGAGAGAGATGGACCTAT 120
Db 61 CAGTCGGGTATCAACTAGGAACACCGAATCTGTATCAAGAGAGAGATGGACCTAT 120
Qy 121 CAGTGGATTTGTCMAACCCAAATCCGTCAACTACAAATTTAAACCAACCCATGTTCCA 180
Db 121 CAGTGGATTTGTCMAACCCAAATCCGTCAACTACAAATTTAAACCAACCCATGTTCCA 180
Qy 181 AAATGGGGGTATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCTCCCGGTGT 240
Db 181 AAATGGGGGTATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCTCCCGGTGT 240
Qy 241 GTTATGTGCTAACAGAGAGACATTTCAATGGGCTAACAAAGAGACAAAGTTCAACAGCAAT 300
Db 241 GTTATGTGCTAACAGAGAGAGACATTTCAATGGGCTAACAAAGAGACAAAGTTCAACAGCAAT 300
Qy 301 TACTTTGGCTCCCTCAGCCAGCCTCAGCTATTCGAGTTGATTCCTTCCAGGAGAGAGAA 360
Db 301 TACTTTGGCTCCCTCAGCCAGCCTCAGCTATTCGAGTTGATTCCTTCCAGGAGAGAGAA 360
Qy 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAAATGTTAATCTTGAAGACATTTGTTGGG 420
Db 361 ATCTATGCCCATTTCAAGAGTCTGCTTCCAAATGTTAATCTTGAAGACATTTGTTGGG 420
Qy 421 GGATGGGATATAGAGCAACATGAACCTGCTGATGTCATGGCCAGAGGCAAAAGGTTTGAAC 480
Db 421 GGATGGGATATAGAGCAACATGAACCTGCTGATGTCATGGCCAGAGGCAAAAGGTTTGAAC 480
Qy 481 ATCGATTTTGCAGAAAGCAGTTGAGGCGCTTACATGGAATCCATGGTTCCACTCCCGGAATC 540
Db 481 ATCGATTTTGCAGAAAGCAGTTGAGGCGCTTACATGGAATCCATGGTTCCACTCCCGGAATC 540
Qy 541 TACGACCCCGGATTTTCACTGCTGCCAACAGAGAGCGTCCAAACAAGTATTAAGGCGC 600
Db 541 TACGACCCCGGATTTTCACTGCTGCCAACAGAGAGCGTCCAAACAAGTATTAAGGCGC 600
Qy 601 ACAAAGCAGAGCAAGTTGAGCAAAATCATCAAGACATCAAGGCGTTTAAGAAAGCACCC 660
Db 601 ACAAAGCAGAGCAAGTTGAGCAAAATCATCAAGACATCAAGGCGTTTAAGAAAGCACCC 660
Qy 661 AAAGTGAACAAGGTGTTGCTGCTGAGCTGCCAACAGAGAGGTATAGCAATTTGGTT 720
Db 661 AAAGTGAACAAGGTGTTGCTGCTGAGCTGCCAACAGAGAGGTATAGCAATTTGGTT 720
Qy 721 GTAGGCGCTTAATGACCACTGAGGAATCTTGGCTGCTGTGAGCAAGAAATGAGCGTGA 780
Db 721 GTAGGCGCTTAATGACCACTGAGGAATCTTGGCTGCTGTGAGCAAGAAATGAGCGTGA 780
Qy 781 ATTTCTCTTCCACCTTGTATGCAATTTGCTGTGATGGAATAATGTTCTTTCAATTAAT 840
Db 781 ATTTCTCTTCCACCTTGTATGCAATTTGCTGTGATGGAATAATGTTCTTTCAATTAAT 840
Qy 841 GGAAGCCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAAACACT 900
Db 841 GGAAGCCCTCAGAAACCTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAAACACT 900
Qy 901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGGTGAAT 960
Db 901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGGTGAAT 960
Qy 961 TTTCTTGTGGGGCTGGTATCAAGCAACATCTATAGTTAGTTACAACCATCTGGGAAC 1020
Db 961 TTTCTTGTGGGGCTGGTATCAAGCAACATCTATAGTTAGTTACAACCATCTGGGAAC 1020
Qy 1021 AATGATGATATGATCTTCGGCTCCAAACCTTCGCTCCAAAGAAATCTTCCAGAGAC 1080
Db 1021 AATGATGATATGATCTTCGGCTCCAAACCTTCGCTCCAAAGAAATCTTCCAGAGAC 1080
Qy 1081 AACGTTTGGAGATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGTAACATCCC 1140
Db 1081 AACGTTTGGAGATATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGTAACATCCC 1140
```

```
Qy 1141 GACCATGTTGTTGTTAATTAAGTATGTGCTTACGTAGGGGATAGCAAGAGCCATGGAT 1200
Db 1141 GACCATGTTGTTGTTAATTAAGTATGTGCTTACGTAGGGGATAGCAAGAGCCATGGAT 1200
Qy 1201 GAGTACATCTTCAAGATATTTATAGGGTGGAAAGAACCAATCTGTTTGGCAACAACATGT 1260
Db 1201 GAGTACATCTTCAAGATATTTATAGGGTGGAAAGAACCAATCTGTTTGGCAACAACATGT 1260
Qy 1261 GAGGATTCCTTTTAAAGTGTGCTCTATTTATCTTGAACCTTGTCCTTGTGAGCTGAGC 1320
Db 1261 GAGGATTCCTTTTAAAGTGTGCTCTATTTATCTTGAACCTTGTCCTTGTGAGCTGAGC 1320
Qy 1321 ACTAGAAATCCAGTTTAAAGCTGAAATAGGGGAAATTCATCTCATTCACCCAGTTGCT 1380
Db 1321 ACTAGAAATCCAGTTTAAAGCTGAAATAGGGGAAATTCATCTCATTCACCCAGTTGCT 1380
Qy 1381 ACCATTTCTACGTATCTTACCAAGGCTCTGTTTCCACCGGGATACCAAGTGTGAT 1440
Db 1381 ACCATTTCTACGTATCTTACCAAGGCTCTGTTTCCACCGGGATACCAAGTGTGAT 1440
Qy 1441 GCATTGTCAAGAGCAGCTGCAATGCTGAGGAAACATATAGAGGCTTGTGATTTGGCC 1500
Db 1441 GCATTGTCAAGAGCAGCTGCAATGCTGAGGAAACATATAGAGGCTTGTGATTTGGCC 1500
Qy 1501 CCAGAAATTAACATGATTTCTGAGTACAAATGA 1533
Db 1501 CCAGAAATTAACATGATTTCTGAGTACAAATGA 1533

RESULT 5
ADS82005
ID ADS82005 standard; cDNA; 1533 BP.
XX
AC ADS82005;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 2.
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; cultivar Wye.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag=a
FT /product="myo-inositol 1-phosphate synthase"
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
PF 11-MAR-2002; 2002US-00025003.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
XX
PI Hitz WD, Sebastian SA;
XX
DR WPI: 2004-639957/62.
DR P-PSDB; ADS82006.
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT phytic acid and inorganic phosphate content of soybean seeds.
XX
PS Claim 2; SEQ ID NO 15; 34pp; English.
XX
```

The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype with (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.

Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other:

Query Match 99.9%; Score 1531.4; DB 13; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGTCATGAGAAATTTAAGTAGAGAGCTTAATGTAAGTACACCGAGACTGAGATT 60
DB 1 ATGTCATGAGAAATTTAAGTAGAGAGCTTAATGTAAGTACACCGAGACTGAGATT 60
QY 61 CAGTCCGTGTACAACTACGAAACCACTGTTCAAGAAACGAGATGGACCTAT 120
DB 61 CAGTCCGTGTACAACTACGAAACCACTGTTCAAGAAACGAGATGGACCTAT 120
QY 121 CAGTGGATTGTCAAAACCAATCCGTCACTACCAATTTAAACCAACCACTGTTCA 180
DB 121 CAGTGGATTGTCAAAACCAATCCGTCACTACCAATTTAAACCAACCACTGTTCA 180
QY 181 AATTTGGGGGTGATGCTTTGGGGTGGGGTGAACAAGGCTTACCTCACCGGTGAT 240
DB 181 AATTTGGGGGTGATGCTTTGGGGTGGGGTGAACAAGGCTTACCTCACCGGTGAT 240
QY 241 GTTATGCTTACAGAGAGGACATTTTCAAGGCTTACAAAGAAAGATTAACAAGCAAT 300
DB 241 GTTATGCTTACAGAGAGGACATTTTCAAGGCTTACAAAGAAAGATTAACAAGCAAT 300
QY 301 TACTTTGCTCCCTCAACCAAGCTCAGCTATTCAGATTGATCTTTCCAGGAGAGAA 360
DB 301 TACTTTGCTCCCTCAACCAAGCTCAGCTATTCAGATTGATCTTTCCAGGAGAGAA 360
QY 361 ATCTATGCCCCATTCAGAGAGTCTGCTTCAATGTTAATCTCGACCACTTGTGTTGG 420
DB 361 ATCTATGCCCCATTCAGAGAGTCTGCTTCAATGTTAATCTCGACCACTTGTGTTGG 420
QY 421 GGATGGATATCGAACAATGAACCTGGCTGATGGCATGGCCAGGCAAAAGGTTTGAC 480
DB 421 GGATGGATATCGAACAATGAACCTGGCTGATGGCATGGCCAGGCAAAAGGTTTGAC 480
QY 481 ATCGATTTTGCAGAAAGAGTTGAGGCTTTACATGAATTCATGGTTCCATCCCGGAAATC 540
DB 481 ATCGATTTTGCAGAAAGAGTTGAGGCTTTACATGAATTCATGGTTCCATCCCGGAAATC 540

```

```

QY 541 TAGACCCGGATTTTCAATGCTGCAACCAAGAGAGCGTCCCAACCAAGTAAAGGCG 600
DB 541 TAGACCCGGATTTTCAATGCTGCAACCAAGAGAGCGTCCCAACCAAGTAAAGGCG 600
QY 601 ACAAAAGCAAGACAGTTTCAGCAAAATCATCAAGACATCAAGGCTTTTAAAGAACCA 660
DB 601 ACAAAAGCAAGACAGTTTCAGCAAAATCATCAAGACATCAAGGCTTTTAAAGAACCA 660
QY 661 AAAGTGACAAAGTGTTGCTCTGAGACTGCAACACAGAGATGATGCAATTTGTT 720
DB 661 AAAGTGACAAAGTGTTGCTCTGAGACTGCAACACAGAGATGATGCAATTTGTT 720
QY 721 GTAGGCTTAATGACCAATGAGATCTTGGCTGCTGAGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTAATGACCAATGAGATCTTGGCTGCTGAGACAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGATGATCCATGCTGCTGATGAGAAATGTTCTTTCAATTAT 840
DB 781 ATTTCTCTTCCACCTTGATGATCCATGCTGCTGATGAGAAATGTTCTTTCAATTAT 840
QY 841 GGAAGCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAACT 900
DB 841 GGAAGCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAACT 900
QY 901 TTGATTTGGTGAATGATCTTCAAGAGTGTGACCAAAATGAAATCTGTGTTGTTAT 960
DB 901 TTGATTTGGTGAATGATCTTCAAGAGTGTGACCAAAATGAAATCTGTGTTGTTAT 960
QY 961 TTTCTTGGGGGCTGATCAAGCAACATCTATAGTTAGTAAACCATCTGGGAAAC 1020
DB 961 TTTCTTGGGGGCTGATCAAGCAACATCTATAGTTAGTAAACCATCTGGGAAAC 1020
QY 1021 AATGATGATGATATCTCGGCTCCAAACCTTCGCTCAAGGAAATCTCCAAAGAC 1080
DB 1021 AATGATGATGATATCTCGGCTCCAAACCTTCGCTCAAGGAAATCTCCAAAGAC 1080
QY 1081 AAGCTTTGGAAGATATGTTCAACAGATGCAATCTCTATGAGACCTGTGTAACATCC 1140
DB 1081 AAGCTTTGGAAGATATGTTCAACAGATGCAATCTCTATGAGACCTGTGTAACATCC 1140
QY 1141 GACCATGTTGTTATTAATGATGCTTACGTAGGAGATAGCAAGACCATGAT 1200
DB 1141 GACCATGTTGTTATTAATGATGCTTACGTAGGAGATAGCAAGACCATGAT 1200
QY 1201 GAGTACACTTCAGAGATATTCATGGGTGAAGAACACATGTTTGGCAACAACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGAAGAACACATGTTTGGCAACAACATGT 1260
QY 1261 GAGGATCCCTTTTACGCTCTCTATTTATCTTGAATGTTGCTTCTGCTGAGCTGAGC 1320
DB 1261 GAGGATCCCTTTTACGCTCTCTATTTATCTTGAATGTTGCTTCTGCTGAGCTGAGC 1320
QY 1321 ACTAGATTCAGATTTAAAGCTGAAATGAGGAAATTCATCTCAATCCACAGTGTCT 1380
DB 1321 ACTAGATTCAGATTTAAAGCTGAAATGAGGAAATTCATCTCAATCCACAGTGTCT 1380
QY 1381 ACCATTTCAAGTATCTGACCAAGGCTCTCTGTTTCCACCGGATACCAAGTGTGAT 1440
DB 1381 ACCATTTCAAGTATCTGACCAAGGCTCTCTGTTTCCACCGGATACCAAGTGTGAT 1440
QY 1441 GCAATGTCAAAAGCAGCGTGAATGCTGAAAAACATTAAGAGGCTTGTGTTGGCC 1500
DB 1441 GCAATGTCAAAAGCAGCGTGAATGCTGAAAAACATTAAGAGGCTTGTGTTGGCC 1500
QY 1501 CCAGAGAAATACATGATCTCGAGTACAAAGTGA 1533
DB 1501 CCAGAGAAATACATGATCTCGAGTACAAAGTGA 1533

```

RESULT 6
AD582003
ID AD582003 standard; cdNA; 1533 BP.
XX

AC ADS82003;
XX
DT 18-NOV-2004 (first entry)
XX Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 4.
XX
XX
XX Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
XX Glycine max; line 29018DP03.
XX
XX Key Location/Qualifiers
XX CDS 1..1533
XX FT /product= "myo-inositol 1-phosphate synthase"
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX MPI; 2004-639957/62.
XX P-PSDB; ADS82004.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Example 8; SEQ ID NO 13; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.
XX
XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
XX
XX Query Match 99.9%; Score 1531.4; DB 13; Length 1533;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGTCATCGGAATTTTAAGTAGAGAGCTCTATGTGAAGTACACCGAGACTGAGATT 60
DB 1 ATGTCATCGGAATTTTAAGTAGAGAGCTCTATGTGAAGTACACCGAGACTGAGATT 60
QY 61 CAGTCGCGTACCACTACGAAACCCAGCACTGTGTCAGAAACGGAATGGACCTAT 120
DB 61 CAGTCGCGTACCACTACGAAACCCAGCACTGTGTCAGAAACGGAATGGACCTAT 120
QY 121 CAGTGGATTGTCAAAACCCAAATCCGTCACTACCAATTTAAACCAACCATGTTCCA 180
DB 121 CAGTGGATTGTCAAAACCCAAATCCGTCACTACCAATTTAAACCAACCATGTTCCA 180
QY 181 AAATGGGGGGTGAATGCTTGTGGGTTGGGGTGAACCAACGGCTTACCGCTGAGT 240
DB 181 AAATGGGGGGTGAATGCTTGTGGGTTGGGGTGAACCAACGGCTTACCGCTGAGT 240
QY 241 GTTATTGCTAACAGAGAGCACTTTCATGGGCTACAAAGACAAAGATTCAACAGCCAAAT 300
DB 241 GTTATTGCTAACAGAGAGCACTTTCATGGGCTACAAAGACAAAGATTCAACAGCCAAAT 300
QY 301 TACTTTGGCTCCCTACCCCAAGCTCAGCTATTCGAGTTGATCCTTCCAGGAGAGAA 360
DB 301 TACTTTGGCTCCCTACCCCAAGCTCAGCTATTCGAGTTGATCCTTCCAGGAGAGAA 360
QY 361 ATCTATGCCCACTTCAAGAGCTGCTTCAATGTTAATCCGACGACATTTGTGTTGG 420
DB 361 ATCTATGCCCACTTCAAGAGCTGCTTCAATGTTAATCCGACGACATTTGTGTTGG 420
QY 421 GGATGGAGATATCAGCAACATGAACTGCTGATGTCATGCGCAGGCAAAAGGTGTTGAC 480
DB 421 GGATGGAGATATCAGCAACATGAACTGCTGATGTCATGCGCAGGCAAAAGGTGTTGAC 480
QY 481 ATCGATTGTCAGACAGTGAAGGCTTATCATGAAATCATGTTCCACTCCCGGAATC 540
DB 481 ATCGATTGTCAGACAGTGAAGGCTTATCATGAAATCATGTTCCACTCCCGGAATC 540
QY 541 TACGACCCGGATTTTCAATGCTGCTCAACCAAGAGAGCGTCCAAACATGTTAAGGCG 600
DB 541 TACGACCCGGATTTTCAATGCTGCTCAACCAAGAGAGCGTCCAAACATGTTAAGGCG 600
QY 601 ACAAAAGCAAGACCAAGTTCAGCAAAATCATCAAAAGCATCAAGCGCTTTAAGGAAGCCACC 660
DB 601 ACAAAAGCAAGACCAAGTTCAGCAAAATCATCAAAAGCATCAAGCGCTTTAAGGAAGCCACC 660
QY 661 AAAGTGACAGAGTGTGTTCTCTGTGACTGCCAACACAGAGATATGCAATTTGGTT 720
DB 661 AAAGTGACAGAGTGTGTTCTCTGTGACTGCCAACACAGAGATATGCAATTTGGTT 720
QY 721 GTAGGCTTAATGACACCATGAGAAATCTTGGGCTGTGAGACGAAATGAGGCTGAG 780
DB 721 GTAGGCTTAATGACACCATGAGAAATCTTGGGCTGTGAGACGAAATGAGGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGTGATGGAATAATGTTCTTTCAATAT 840
DB 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGTGATGGAATAATGTTCTTTCAATAT 840
QY 841 GGAAGCCCTCAGAACCTTTTGTACAGGGCTGATTTGATCTTGGCCATCCGAGGAACACT 900
DB 841 GGAAGCCCTCAGAACCTTTTGTACAGGGCTGATTTGATCTTGGCCATCCGAGGAACACT 900
QY 901 TTGATTGGTGAAGTACTTCAAGAGTGTGACCAAAATGAATACTGTGTTGTTGAT 960
DB 901 TTGATTGGTGAAGTACTTCAAGAGTGTGACCAAAATGAATACTGTGTTGTTGAT 960
QY 961 TTTCTTGTGGGGCTGTGTATCAAGCCACATCTATAGTTAGTACCAACCATCTGGGAAAC 1020
DB 961 TTTCTTGTGGGGCTGTGTATCAAGCCACATCTATAGTTAGTACCAACCATCTGGGAAAC 1020
QY 1021 AATGATGTATGATATCTCGGCTCCAAACCTTCCGCTCCAGGAATCTTCCAAAGAC 1080
DB 1021 AATGATGTATGATATCTCGGCTCCAAACCTTCCGCTCCAGGAATCTTCCAAAGAC 1080

Db 1021 AATGATGTAATGATCTCTGGGCTCCAAACCTTCGCTCCAAAGAAATCTCCAAAGG 1080
 Qy 1081 AACCTTGTGACATATGCTGCAACAGCAATGCAATCTCTATATAGCTGTGAACTCC 1140
 Db 1081 AACCTTGTGACATATGCTGCAACAGCAATGCAATCTCTATATAGCTGTGAACTCC 1140
 Qy 1141 GACCATGTGTTGTTATTAAGTATGCTTACGTAGGGGATAGCAAGAGCATGGAT 1200
 Db 1141 GACCATGTGTTGTTATTAAGTATGCTTACGTAGGGGATAGCAAGAGCATGGAT 1200
 Qy 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCATTTGTCACACATGT 1260
 Db 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCATTTGTCACACATGT 1260
 Qy 1261 GAGGATTCCTTTTACTGCTCTATATATCTTGAATGCTTCTTGTGCTGAGCTGAGC 1320
 Db 1261 GAGGATTCCTTTTACTGCTCTATATATCTTGAATGCTTCTTGTGCTGAGCTGAGC 1320
 Qy 1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGAAATTCGACTCATTTCCACCGAGTTGCT 1380
 Db 1321 ACTAGAAATCCAGTTTAAAGCTGAAATGAGGAAATTCGACTCATTTCCACCGAGTTGCT 1380
 Qy 1381 ACCATTCTCAGCTATCTGACCAAGGCTCCTTGTGTTCCACCGGATACACCATGTGTAAT 1440
 Db 1381 ACCATTCTCAGCTATCTGACCAAGGCTCCTTGTGTTCCACCGGATACACCATGTGTAAT 1440
 Qy 1441 GCATTGTCAAACGAGCTGCAATGCTGGAAGAAACATATATAGGGCTTGTGATTTGGCC 1500
 Db 1441 GCATTGTCAAACGAGCTGCAATGCTGGAAGAAACATATATAGGGCTTGTGATTTGGCC 1500
 Qy 1501 CCAGAGAAATACATGATCTCTCGAGTACAGTGA 1533
 Db 1501 CCAGAGAAATACATGATCTCTCGAGTACAGTGA 1533

RESULT 7
 ADQ14498
 ID ADQ14498 standard; cDNA; 1533 BP.

XX ADQ14498;

DT 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #2.

KM Soybean; myo-inositol 1-phosphate synthase; gene; ss;
 KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
 KM inorganic phosphate; mutant.

XX Glycine max.

OS Synthetic.

FH Key Location/Qualifiers
 FT 1. 1533
 FT CDS /tag= a
 FT /product= "mutant soybean myo-inositol 1-phosphate
 FT synthase #2"

PN US2004128713-A1.

PD 01-JUL-2004.

PF 21-NOV-2003; 2003US-00718955.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1998; 98WO-US006822.

PR 26-APR-1999; 99US-00299315.

PR 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.

PA (STRE/) STREIT L G.

XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
 PI MPI; 2004-533135/51.
 XX P-PSDB; ADQ14499.
 DR
 DR
 PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytic acid and increased sucrose, leading to valuable and useful soybean
 PT products.

Example 8; SEQ ID NO 9; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
 CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
 CC phosphate. The invention also relates to a chimeric gene operably linked
 CC to suitable regulatory sequences, where expression of the chimeric gene
 CC results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
 CC comprising the chimeric gene, a method of making the soybean plant, a
 CC seed of the soybean plant, a soy protein product derived from the
 CC processing of soybean seeds, a method of making or producing a soy
 CC protein product and a method of using a soybean plant homozygous for at
 CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 CC nucleic acid fragment and methods are useful for producing plants with
 CC decreased raffinose, stachyose and phytic acid content and increased
 CC sucrose and inorganic phosphate content, leading to valuable and useful
 CC soybean products. This sequence represents cDNA encoding a mutant soybean
 CC myo-inositol 1-phosphate synthase polypeptide of the invention.

Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Query Match 95.5%; Score 1464.2; DB 12; Length 1533;
 Best Local Similarity 97.2%; Pred. No. 0;
 Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 1 ATGTTCAATCAGAAATTTTAAAGTAGAGAGCTCTTAATGTGAATACACCGAGACTGAGATT 60
 Db 1 ATGTTCAATCAGAAATTTTAAAGTAGAGAGCTCTTAATGTGAATACACCGAGACTGAGATT 60
 Qy 61 CAGTCGGGTGTAACATACGAAACCAACCACTTTGTTCCAGAAACAGAAATGGCACTTAT 120
 Db 61 CAGTCGGGTGTAACATACGAAACCAACCACTTTGTTCCAGAAACAGAAATGGCACTTAT 120
 Qy 121 CAGTGAATTTGTAACCAACCAATCCGTCAACTACATTTTAAACCAACCAACCATGTTCCA 180
 Db 121 CAGTGAATTTGTAACCAACCAATCCGTCAACTACATTTTAAACCAACCAACCATGTTCTT 180
 Qy 181 AAATTGGGGGTGATGCTTGTGGGTGGGTTGAAACCAACCGCTTACCCGTCAACCGGTGT 240
 Db 181 AAATTGGGGGTGATGCTTGTGGGTGGGTTGAAACCAACCGCTTACCCGTCAACCGGTGT 240
 Qy 241 GTTATTTGTAACAGAGAGACATTTTCAATGAGGCTTACAAAGAGACAAATTTCAACACCAAT 300
 Db 241 GTTATTTGTAACAGAGAGACATTTTCAATGAGGCTTACAAAGAGACAAATTTCAACACCAAT 300
 Qy 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATCTCTTCCAGGAGAGAGAA 360
 Db 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGATCTCTTCCAGGAGAGAGAA 360
 Qy 361 ATCTATGCCCCATTCAGAGAGCTGCTTCAATGTTTAACTCTGACAGCATTTGTGTTGGG 420
 Db 361 ATCTATGCCCCATTCAGAGAGCTGCTTCAATGTTTAACTCTGACAGCATTTGTGTTGGG 420
 Qy 421 GAGTGGATATCAGAAATATGATCTGCTGATGCAATGAGGCAAGGCAAGGTGTTTAC 480
 Db 421 GAGTGGATATCAGAAATATGATCTGCTGATGCAATGAGGCAAGGCAAGGTGTTTAC 480
 Qy 481 ATCGATTTGCAAGAGCACTTGAAGGCTTACATGGAATCATGATGTTCCACTCCCGGAATC 540
 Db 481 ATCGATTTGCAAGAGCACTTGAAGGCTTACATGGAATCATGATGTTCCACTCCCGGAATC 540


```

QY 541 TACGACCCGCGATTTCATGCTGCGCAACCAAGAGAGCGTCCACAACGCTGATTAAAGGCG 600
DB 541 TATGACCCGCGATTTCATGCTGCGCAACCAAGAGAGCGTCCACAACGCTGATTAAAGGCG 600
QY 601 ACAAAGCAAGACCAAGTTCAGCAATCATCAAAAGACATCAAGCCGTTTAAAGAACCCACC 660
DB 601 ACAAAGCAAGACCAAGTTCAGCAATCATCAAAAGACATCAAGCCGTTTAAAGAACCCACC 660
QY 661 AAAGTGAACAAGTGGTGTCCGTGAGACTGCGCAACAGAGAGGATATGCAATTTGGTT 720
DB 661 AAAGTGAACAAGTGGTGTCTGAGACTGCGCAACAGAGAGGATATGCAATTTGGTT 720
QY 721 GTAGGCTTTAATGACACCATGAGAAATCTTTGGCTGTGTGAGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTTAATGACACCATGAGAAATCTTTGGCTGTGTGAGACAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTTCCACCTTTGATGCGCATTCGCTGTGTGATGAGAAATGTTCTTTCAATTAAT 840
DB 781 ATTTCTCTTCCACCTTTGATGCGCATTCGCTGTGTGATGAGAAATGTTCTTTCAATTAAT 840
QY 841 GGAAGCCCTCAGAAACCTTTTGNACAGGGGCTGATTGATCTTGCCATCGCGAGGAACACT 900
DB 841 GGAAGCCCTCAGAAACCTTTTGNACAGGGGCTGATTGATCTTGCCATCGCGAGGAACACT 900
QY 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTGTTGAT 960
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTGTTGAT 960
QY 961 TTTCTTTGTGGGGGTGTGATCAAGCCCAACATCTATAGTTAACAACCATCTGGGAAC 1020
DB 961 TTTCTTTGTGGGGGTGTGATCAAGCCCAACATCTATAGTTAACAACCATCTGGGAAC 1020
QY 1021 AATGATGATGATGATCTCTCGGCTCCACAACCTTCGCTCCAAAGGAATCTTCCAAGAC 1080
DB 1021 AATGATGATGATGATCTCTCGGCTCCACAACCTTCGCTCCAAAGGAATCTTCCAAGAC 1080
QY 1081 AACGTTGTGAGATATGCTGCAACAGCAATGCGCATCTCTATGAGCCTTGGAACATCCC 1140
DB 1081 AACGTTGTGAGATATGCTGCAACAGCAATGCGCATCTCTATGAGCCTTGGAACATCCA 1140
QY 1141 GACCATGTTGTTTATTAAGTATGCTTACGTAAGGGGATAGCAAGAGCCATGAT 1200
DB 1141 GACCATGTTGTTTATTAAGTATGCTTACGTAAGGGGATAGCAAGAGCCATGAT 1200
QY 1201 GAGTACACTTCAGAGATTTCAATGAGGTGGAAGAACACCATGTTTGGACAACACATGT 1260
DB 1201 GAGTACACTTCAGAGATTTCAATGAGGTGGAAGAACACCATGTTTGGACAACACATGT 1260
QY 1261 GAGGATTCCTTTTAACTGCTCTATTAATCTTGAATGCTGCTTCTTGTGAGCTGAGC 1320
DB 1261 GAGGATTCCTTTTAACTGCTCTATTAATCTTGAATGCTGCTTCTTGTGAGCTGAGC 1320
QY 1321 ACTAGATTCAGTTTAAAGTGAAGGGAATTCACACTCATTCCACCAAGTGTGCT 1380
DB 1321 ACTAGATTCAGTTTAAAGTGAAGGGAATTCACACTCATTCCACCAAGTGTGCT 1380
QY 1381 ACCATCTCAGCTATGACCAAGGCTCCTGCTGCTCAACCGGATCACACAGTGTGAAT 1440
DB 1381 ACCATCTCAGCTATGACCAAGGCTCCTGCTGCTCAACCGGATCACACAGTGTGAAT 1440
QY 1441 GGAATGTCAAAGAGCGTGCATGCTGGAACCATTAATGAGGCTTGTGTGATTGGCC 1500
DB 1441 GGAATGTCAAAGAGCGTGCATGCTGGAACCATTAATGAGGCTTGTGTGATTGGCC 1500
QY 1501 CCAGAGAAATAACTATGTTCTCGAGTCAAGTGA 1533
DB 1501 CCAGAGAAATAACTATGTTCTCGAGTCAAGTGA 1533

```

RESULT 8
ADS81999
ID ADS81999 standard; cDNA; 1533 BP.

```

XX AC ADS81999;
XX XX 18-NOV-2004 (first entry)
DT DT
XX XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 3.
XX KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX KW raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX OS Glycine max; line 29004UP01.
XX FH
XX FT Key Location/Qualifiers
XX FT CDS 1..1533
XX FT /*tag= a
XX FT /product= "myo-inositol 1-phosphate synthase"
XX XX
XX PN US2003074685-A1.
XX XX
XX PD 17-APR-2003.
XX XX
XX PF 11-MAR-2002; 2002US-00025003.
XX XX
XX PR 08-APR-1997; 97US-00835751.
XX PR 07-APR-1996; 96WO-US006822.
XX PA (HITZ/) HITZ W D.
XX PA (SEBA/) SEBASTIAN S A.
XX XX
XX PI Hitz WD, Sebastian SA;
XX XX
XX DR WPI; 2004-639957/62.
XX DR P-PsDB; ADS82000.
XX XX
XX PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX PT phytic acid and inorganic phosphate content of soybean seeds.
XX PS
XX PS Example 8; SEQ ID NO 9; 34p; English.
XX XX
XX CC The invention relates to an isolated nucleic acid fragment encoding a
XX CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX CC phosphate synthase having decreasing capacity for the synthesis for myo-
XX CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX CC complement, subfragment or the complement of the subfragment, operably
XX CC linked to suitable regulatory sequences, where expression of the chimeric
XX CC gene results in a decrease in expression of an endogenous or native gene
XX CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
XX CC acid content of less than 17 micromol/g, a seed content of raffinose plus
XX CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
XX CC from the plant, making a soybean plant with the heritable phenotype
XX CC (comprising crossing LR33 or the plant comprising the chimeric gene with
XX CC an elite soybean plant and selecting a progeny plant of the cross of
XX CC crossing step that has a heritable phenotype as mentioned above), seeds
XX CC of soybean plant made by the above method, a soy protein product derived
XX CC from seeds of a soybean plant (homozygous for one or more gene encoding a
XX CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
XX CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
XX CC heritable phenotype as mentioned above), and making or producing a
XX CC soybean protein product derived from seeds of a soybean plant with
XX CC heritable phenotype as mentioned above. The nucleic acid is useful for
XX CC altering raffinose saccharide, sucrose, phytic acid and inorganic
XX CC phosphate content of soybean seeds thus leading to valuable and useful
XX CC soybean products, since the presence of high concentration of raffinose
XX CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
XX CC when consumed by humans. The present sequence encodes a wild-type myo-
XX CC inositol 1-phosphate synthase.
XX XX
XX SQ Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

```


Query Match 95.5%; Score 1464.2; DB 13; Length 1533;
 Best Local Similarity 97.2%; Pred. No. 0;
 Matches 1450; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```

QY 1 ATGTCATCGAGAAATTTTAAGTAGAGAGTCTTAATGTGAATGACCCGAGACTGAGATT 60
DB 1 ATGTCATCGAGAAATTTTAAGTAGAGTCTTAATGTGAATGACCCGAGACTGAGATT 60
QY 61 CAGTCCGTGTACACATACGAAACACCGAACTTGTTCAAGAAACAGAAATGCGACTAT 120
DB 61 CAGTCCGTGTACACATACGAAACACCGAACTTGTTCAAGAAACAGAAATGCGACTAT 120
QY 121 CAGTGAATGTCAAAACCAATCGTCAACTACCAATTTTAAACCAACCCATGTTTCA 180
DB 121 CAGTGAATGTCAAAACCAATCGTCAACTACCAATTTTAAACCAACCCATGTTTCT 180
QY 181 AAATGGGGGTGATGCTGTGGGTGGGTGGAACAACGGCTCTACCCCTCACGGTGGT 240
DB 181 AAATGGGGGTGATGCTGTGGGTGGGTGGAACAACGGCTCTACCCCTCACGGTGGT 240
QY 241 GTTATTGCTTAACAGAGACATTTATGAGGCTTAAAGAACAAAGATTCAACAGCCAA 300
DB 241 GTTATTGCTTAACAGAGACATTTATGAGGCTTAAAGAACAAAGATTCAACAGCCAA 300
QY 301 TACTTGGCTCCCTCAACCAAGCTGAGCTATTGAGTTGAGTCTTCCAGGGAGAGAA 360
DB 301 TACTTGGCTCCCTCAACCAAGCTGAGCTATTGAGTTGAGTCTTCCAGGGAGAGAA 360
QY 361 ATCTATGCCCATTTCAAGAGTGTCTTCCAAATGGTTAATCTGTGACACATTTGTTGGG 420
DB 361 ATCTATGCCCATTTCAAGAGTGTCTTCCAAATGGTTAATCTGTGACACATTTGTTGGG 420
QY 421 GGATGGATATACGAAATGAACTGGCTGATGCAATGGCCAGGGCAAAAGTGTGAC 480
DB 421 GGATGGATATACGAAATGAACTGGCTGATGCAATGGCCAGGGCAAAAGTGTGAC 480
QY 481 ATCGATTTGACAGAGAGTTGAGGCTTTCATGGAATTCATGCTTCCACCCCGGAATC 540
DB 481 ATCGATTTGACAGAGAGTTGAGGCTTTCATGGAATTCATGCTTCCACCCCGGAATC 540
QY 541 TACGACCCGGAATTTCAATGCTGCAACCAAGAGACGTCGCAACCAAGTTAAGGCG 600
DB 541 TACGACCCGGAATTTCAATGCTGCAACCAAGAGACGTCGCAACCAAGTTAAGGCG 600
QY 601 ACAAGACAGAGCAATTCAGCAATCATCAAGACATCAAGGCGTTTAAAGAACCC 660
DB 601 ACAAGACAGAGCAATTCAGCAATCATCAAGACATCAAGGCGTTTAAAGAACCC 660
QY 661 AAAGTGAAGAAGGTGTGCTGCTGAGACCTGCCAACAAGAGGTATAGCAATTTGGTT 720
DB 661 AAAGTGAAGAAGGTGTGCTGCTGAGACCTGCCAACAAGAGGTATAGCAATTTGGTT 720
QY 721 GTAGGCTTAATGACACATGAGAACTCTTGAGCTGTGTGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTAATGACACATGAGAACTCTTGAGCTGTGTGACAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGATGAAATGTTCTTTCAATTAAT 840
DB 781 ATTTCTCTTCCACCTTGTATGCAATGCTGTGTGATGAAATGTTCTTTCAATTAAT 840
QY 841 GGAAGCCCTCAGAAACCTTTTGTACCAAGGCTGATTTGATCTTGCCATCGCGAGAACT 900
DB 841 GGAAGCCCTCAGAAACCTTTTGTACCAAGGCTGATTTGATCTTGCCATCGCGAGAACT 900
QY 901 TTGATTTGAGAGTACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGTTGAT 960
DB 901 TTGATTTGAGAGTACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGTTGAT 960
QY 961 TTTCTTTGGGGGCTGTGATCAAGCAACATCTAATAGTTTCAACCACTTGGGAAC 1020
DB 961 TTTCTTTGGGGGCTGTGATCAAGCAACATCTAATAGTTTCAACCACTTGGGAAC 1020
QY 1021 AATGATGATGATATCTCGGCTCCACAAACCTTCGCTCCAGGAATCTTCCAGAGC 1080
  
```

```

DB 1021 AATGATGATGATATCTTCGCTCCACAAACCTTCGCTCCAGGAATCTTCCAGAGC 1080
QY 1081 AACGTTGTGACGATATGCTCAACAGCAATGCAATCTCTATGAGCTGTGTAACATCC 1140
DB 1081 AACGTTGTGATGATATGCTCAACAGCAATGCAATCTCTATGAGCTGTGTAACATCC 1140
QY 1141 GACCATGTTGTTATTTAAGTATGCTTACGTAGGGGATATGCAAGAGCATGAT 1200
DB 1141 GACCATGTTGTTATTTAAGTATGCTTACGTAGGGGATATGCAAGAGCATGAT 1200
QY 1201 GAGTACCTTCAGAGATATTCATGAGGTGAAAGAACCAATGTTTGTGACAAACATGT 1260
DB 1201 GAGTACCTTCAGAGATATTCATGAGGTGAAAGAACCAATGTTTGTGACAAACATGT 1260
QY 1261 GAGGATTCCTTTTACGCTCTCTATTAATCTTGATCTTCTTCTGAGCTGAGC 1320
DB 1261 GAGGATTCCTCTTACGCTCTCTATTAATCTTGATCTTCTTCTGAGCTGAGC 1320
QY 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACCTCATTTCCACCACTTGT 1380
DB 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACCTCATTTCCACCACTTGT 1380
QY 1381 ACCATTCGAGCTATCTGACCAAGGCTCTGTTTCCACCGGGTACACAGTGTGAAT 1440
DB 1381 ACCATTCGAGCTATCTGACCAAGGCTCTGTTTCCACCGGGTACACAGTGTGAAT 1440
QY 1441 GCATTGTCAAAACAGGCTGCAATGCTGAAACCATTAATGAGGCTTGTGATTTGGC 1500
DB 1441 GCATTGTCAAAACAGGCTGCAATGCTGAAACCATTAATGAGGCTTGTGATTTGGC 1500
QY 1501 CCAGAGATTAACATGATTTCTGAGTACAAATGA 1533
DB 1501 CCAGAGATTAACATGATTTCTGAGTACAAATGA 1533
  
```

RESULT 9
 ADS81993
 ID ADS81993 standard; cDNA; 1533 BP.
 XX
 AC ADS81993;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 1.
 XX
 KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
 XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
 OS Glycine max; cultivar Wye.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1533
 FT /tag= a
 FT /product= "myo-inositol 1-phosphate synthase"
 XX
 PN US2003074685-A1.
 PD 17-APR-2003.
 XX
 PF 11-MAR-2002; 2002US-00025003.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 XX
 PI Hitz WD, Sebastian SA;
 XX
 DR MPI; 2004-639957/62.
 DR P-PsDB; ADS81994.
 XX

PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

PS Claim 2; SEQ ID NO 1; 34bp; English.

XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.

CC Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

QQ Query Match 95.5%; Score 1464.2; DB 13; Length 1533;

Beet Local Similarity 97.2%; Pred. No. 0;

Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 ATGTCATCGAGAAATTTAAGTAGAGAGTCTTAATGTAGATACACCGAGACTGAGATT 60
DB 1 ATGTCATCGAGAAATTTAAGTAGAGTGTCTTAATGTAGATACACCGAGACTGAGATT 60
QY 61 CAGTCCGTGTACACTACGAAACACCGAATTGTTACGAGAACGGAATGGCACTAT 120
DB 61 CAGTCCGTGTACACTACGAAACACCGAATTGTTACGAGAACGGAATGGCACTAT 120
QY 121 CAGTGGATGTGAACCCAAATCCGTCATCAATTTAAACCAACGCCATGTTCCA 180
DB 121 CAGTGGATGTGAACCCAAATCCGTCATCAATTTAAACCAACGCCATGTTCCA 180
QY 181 AATTTGGGGGTGAGTCTTGTGGGTTGGGTAACAACGCGCTTCAACCGGTGAT 240
DB 181 AATTTGGGGGTGAGTCTTGTGGGTTGGGTAACAACGCGCTTCAACCGGTGAT 240
QY 241 GTTATATGCTAACCGAAGGGCAATTTATGAGGCTAACAAAGAACAAAGATTCACACGCAAT 300
DB 241 GTTATATGCTAACCGAAGGGCAATTTATGAGGCTAACAAAGAACAAAGATTCACACGCAAT 300
QY 301 TACTTTGGGCTCCCTCCACCAAGCTCAGCTATTCGAGTTGGATCTTCCAGGAGAGGAA 360
DB 301 TACTTTGGGCTCCCTCCACCAAGCTCAGCTATTCGAGTTGGATCTTCCAGGAGAGGAA 360
QY 361 ATCTATGCCCATTTCAAGAGTGTCTTCAATGTATTAATCTGACGACATTTGTTGGG 420
DB 361 ATCTATGCCCATTTCAAGAGTGTCTTCAATGTATTAATCTGACGACATTTGTTGGG 420
QY 421 GGATGGATATACGAACTGAACCTGGCTGATGCTATGCGACGAGCAAGGTGTTGAC 480
DB 421 GGATGGATATACGAACTGAACCTGGCTGATGCTATGCGACGAGCAAGGTGTTGAC 480

DB 421 GGATGGATATACGAACTGAACCTGGCTGATGCTATGCGACGAGCAAGGTGTTGAC 480
QY 481 ATCTATGCCCATTTCAAGAGTGTCTTCAATGTATTAATCTGACGACATTTGTTGGG 540
DB 481 ATCTATGCCCATTTCAAGAGTGTCTTCAATGTATTAATCTGACGACATTTGTTGGG 540
QY 541 TATGACCCCGGATTTCAATGCTGCGCAACCAAGAGGCTGTCCACMAAGTCAACAGGAG 600
DB 541 TATGACCCCGGATTTCAATGCTGCGCAACCAAGAGGCTGTCCACMAAGTCAACAGGAG 600
QY 601 ACAAAGCAAGGCAAGTTCAGCAATATCATCAAGACATCAAGGCTTAAAGAGCCACC 660
DB 601 ACAAAGCAAGGCAAGTTCAGCAATATCATCAAGACATCAAGGCTTAAAGAGCCACC 660
QY 661 AAAGTGCAAGAGTGTGTCTGTGACCTGCCAACAGAGAGGTATAGCAATTTGGTT 720
DB 661 AAAGTGCAAGAGTGTGTGTGACCTGCCAACAGAGAGGTATAGCAATTTGGTT 720
QY 721 GTAGGCTTTAATGACACCATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
DB 721 GTAGGCTTTAATGACACCATGAGAAATCTTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGTATGCAATTCCTGTGTGAGGAAATGTTCTTTCAATAT 840
DB 781 ATTTCTCTTCCACCTTGTATGCAATTCCTGTGTGAGGAAATGTTCTTTCAATAT 840
QY 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTGTATCTTCCATCGCAGGAACACT 900
DB 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTGTATCTTCCATCGCAGGAACACT 900
QY 901 TTGATTTGGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGAT 960
DB 901 TTGATTTGGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGAT 960
QY 961 TTTCTTGGGGGGCTGTATGAAGCAACATCTATAGTTAGTTCAACATCTGGGAAC 1020
DB 961 TTTCTTGGGGGGCTGTATGAAGCAACATCTATAGTTAGTTCAACATCTGGGAAC 1020
QY 1021 AATGATGTATGAATCTCTCGGCTTCCAAACCTTCCGCTCCAGGAATCTCCAAAGAC 1080
DB 1021 AATGATGTATGAATCTCTCGGCTTCCAAACCTTCCGCTCCAGGAATCTCCAAAGAC 1080
QY 1081 AAGTTTGTGAGATATGATGTGCAACAGCAATGCAATCTCTATGAGCTGTGGAACATCCA 1140
DB 1081 AAGTTTGTGAGATATGATGTGCAACAGCAATGCAATCTCTATGAGCTGTGGAACATCCA 1140
QY 1141 GACCAATGTTGTTAATTAAGATGTGCTTACGTAAGGGGATAGCAAGAGCCATGTAT 1200
DB 1141 GACCAATGTTGTTAATTAAGATGTGCTTACGTAAGGGGATAGCAAGAGCCATGTAT 1200
QY 1201 GAGTACACTTCAGAGATATTCATGAGGTGAAAGAACACCATTTGTTTCCACAAACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTCATGAGGTGAAAGAACACCATTTGTTTCCACAAACATGT 1260
QY 1261 GAGATTCCTTTTATGCTGCTCTATATCTTGAACCTTGTCTTCTTGTGAGCTGAGC 1320
DB 1261 GAGATTCCTTTTATGCTGCTCTATATCTTGAACCTTGTCTTCTTGTGAGCTGAGC 1320
QY 1321 ACTAAGATTCAGTTTAAAGCTGAAATAGAGGAAATTCACCTCATCCACCACTGCT 1380
DB 1321 ACTAAGATTCAGTTTAAAGCTGAAATAGAGGAAATTCACCTCATCCACCACTGCT 1380
QY 1381 ACCATTTCCAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGATACCAAGTGTGAT 1440
DB 1381 ACCATTTCCAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGATACCAAGTGTGAT 1440
QY 1441 GCATTTGTCAAAAGCAGCTGCAATGTCTGAAAAAATAGAGGCTTGTGTTGATGGCC 1500
DB 1441 GCATTTGTCAAAAGCAGCTGCAATGTCTGAAAAAATAGAGGCTTGTGTTGATGGCC 1500
QY 1501 CCAAGGAATACATGATTTCTGAGTACAAAGTGA 1533
DB 1501 CCAAGGAATACATGATTTCTGAGTACAAAGTGA 1533

RESULT 10
ADQ14490
ID ADQ14490 standard; cDNA; 1760 BP.
AC ADQ14490;
XX
DT 23-SEP-2004 (first entry)
XX
DE Wild type soybean myo-inositol 1-phosphate synthase cDNA #1.
XX
KM Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KM inorganic phosphate.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 54..1586
FT /tag= a
FT /product= "Wild type soybean myo-inositol 1-phosphate
FT synthase #1"
XX
PN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PF 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WC-US006822.
PR 26-APR-1999; 99US-00299315.
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX
DR MPI: 2004-53335/51.
DR P-PSDB; ADQ14491.
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
PS Claim 4; SEQ ID NO 1; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
XX inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
XX synthase having a decreased capacity for the synthesis of myo-inositol 1-
XX phosphate. The invention also relates to a chimeric gene operably linked
XX to suitable regulatory sequences, where expression of the chimeric gene
XX results in a decrease in expression of an endogenous or native gene
XX encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
XX comprising the chimeric gene, a method of making the soybean plant, a
XX seed of the soybean plant, a soy protein product derived from the
XX processing of soybean seeds, a method of making or producing a soy
XX protein product and a method of using a soybean plant homozygous for at
XX least one gene encoding a mutant myo-inositol 1-phosphate synthase having
XX decreased capacity for the synthesis of myo-inositol 1-phosphate. The
XX nucleic acid fragment and methods are useful for producing plants with
XX decreased raffinose, stachyose and phytic acid content and increased
XX sucrose and inorganic phosphate content, leading to valuable and useful
XX soybean products. This sequence represents cDNA encoding a wild type
XX soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
SQ Sequence 1760 BP; 494 A; 371 C; 411 G; 484 T; 0 U; 0 Other;

Query Match 95.5%; Score 1464.2; DB 12; Length 1760;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 1 ATGTCATGAGAAATTTTAAAGTAGAGAGCTCAATAGTGAAGTCAACCGAGACTGAGATT 60
DB 54 ATGTCATGAGAAATTTTAAAGTAGAGAGCTCAATAGTGAAGTCAACCGAGACTGAGATT 113
QY 61 CAGTCCGTGTACAACTACGAAACCAACGAACTTGTTCACGAGAACGAAATGACCTAT 120
DB 114 CAGTCCGTGTACAACTACGAAACCAACGAACTTGTTCACGAGAACGAAATGACCTAT 173
QY 121 CAGTGAATGTCAAAACCAATCCGTCACTACCAATTTAAACCAACCACTGTTCCA 180
DB 174 CAGTGAATGTCAAAACCAATCCGTCACTACCAATTTAAACCAACCACTGTTCCCT 233
QY 181 AAATGGGGGTGATGCTGTGGGGGTGGGAAACACGGCTCTACCGCTGCTGCTGCTGCT 240
DB 234 AAATGGGGGTGATGCTGTGGGGGTGGGAAACACGGCTCTACCGCTGCTGCTGCTGCT 293
QY 241 GTTATTTCTAACAGAGAGCAATTTTCATGAGGCTACAAAGGACAAAGATTCAACGCAAT 300
DB 294 GTTATTTCTAACAGAGAGCAATTTTCATGAGGCTACAAAGGACAAAGATTCAACGCAAT 353
QY 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTTCAGATTGATCCTTCCAGGAGAGAA 360
DB 354 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTTCAGATTGATCCTTCCAGGAGAGAA 413
QY 361 ATCTATGCCCCATTCGAAGAGTGTGCTTCCAAAGGTTAACTCGAAGCAATTTGTTGGG 420
DB 414 ATCTATGCCCCATTCGAAGAGTGTGCTTCCAAAGGTTAACTCGAAGCAATTTGTTGGG 473
QY 421 GGATGGGATATGCAACATGAACCTGCTGATGTCATGAGGACGAGGCAAAAGGTGTTGAC 480
DB 474 GGATGGGATATGCAACATGAACCTGCTGATGTCATGAGGACGAGGCAAAAGGTGTTGAC 533
QY 481 ATCGATTTCAGAGAGAGTTGAGGCTTATCATGAAATCATGTTCCACTCCCGGAATC 540
DB 534 ATCGATTTCAGAGAGAGTTGAGGCTTATCATGAAATCATGTTCCACTCCCGGAATC 593
QY 541 TACGACCCGGATTTTCATGCTGCTGCCAAACGAAGGAGCGGCCAAACAGCTGTTAAGGCG 600
DB 594 TACGACCCGGATTTTCATGCTGCTGCCAAACGAAGGAGCGGCCAAACAGCTGTTAAGGCG 653
QY 601 ACAAGCAAGACAAAGTTGACGAAATCATCAAGACATCAAGACGCTTAAAGAGGACACC 660
DB 654 ACAAGCAAGACAAAGTTGACGAAATCATCAAGACATCAAGACGCTTAAAGAGGACACC 713
QY 661 AAAGTGACAGAGTGTGCTGCTGAGCTGCCAACACAGAGGTTANAGCAATTTGGTT 720
DB 714 AAAGTGACAGAGTGTGCTGCTGAGCTGCCAACACAGAGGTTANAGCAATTTGGTT 773
QY 721 GTAGGCTTAAAGACACCATGAGAAATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 774 GTAGGCTTAAAGACACCATGAGAAATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 833
QY 781 ATTTCTCTTCCACCTTGTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 834 ATTTCTCTTCCACCTTGTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 893
QY 841 GGAAGCCTCAGACACCTTTGTACCAAGGCTGATTTGATCTTTCATGCGAGAACT 900
DB 894 GGAAGCCTCAGACACCTTTGTACCAAGGCTGATTTGATCTTTCATGCGAGAACT 953
QY 901 TTGATTTGAGAGAGACTTCAAGAGTGTGACCAAAATGAATGCTGTGTTGTTAT 960
DB 954 TTGATTTGAGAGAGACTTCAAGAGTGTGACCAAAATGAATGCTGTGTTGTTAT 1013
QY 961 TTTCTTGTGGGGCTGTATCAAGCAACATCTATAGTTAGTTACCAACATCTGGAAC 1020
DB 1014 TTTCTTGTGGGGCTGTATCAAGCAACATCTATAGTTAGTTACCAACATCTGGAAC 1073
QY 1021 AATGATGATGATATCTCGGCTCCAAACCTTCCGCTCCAAAGAAATCTCCAAAGAC 1080

```
Db 1074 AATGATGATGAATCTTGGGCTCCAAACTTCCGTTCCAGGAAATCTCCAAAGAC 1133
Qy 1081 AACGTTGTTGAGATATGTTCAACAGCAATGCCATCTCTTAAGACCTGTAATCC 1140
Db 1134 AACGTTGTTGATATGTTCAACAGCAATGCCATCTCTTAAGACCTGTAATCC 1193
Qy 1141 GACCATGTTGTTATTAAGATATGCTTTCGTAAGGAGATAGCAAGAGCCATGAT 1200
Db 1194 GACCATGTTGTTATTAAGATATGCTTTCGTAAGGAGATAGCAAGAGCCATGAT 1253
Qy 1201 GAGTCACTTCAAGATATTCATGGGTGGAAGAACACCATTTGTTGCAACATGT 1260
Db 1254 GAGTCACTTCAAGATATTCATGGGTGGAAGAACACCATTTGTTGCAACATGT 1313
Qy 1261 GAGGATCCCTTTGAGCTCTCTTAATCTTGAAGCTGCTCTTGTGAGCTGAGC 1320
Db 1314 GAGGATCCCTTTGAGCTCTCTTAATCTTGAAGCTGCTCTTGTGAGCTGAGC 1373
Qy 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACATTCACCCAGTTGCT 1380
Db 1374 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCACATTCACCCAGTTGCT 1433
Qy 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGTAACACAGTGTGAT 1440
Db 1434 ACCATTTCTAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGTAACACAGTGTGAT 1493
Qy 1441 GATTTCTCAACAGCGTGCATTCGTGGAAGACATTAATAGGGCTTGTGATTTGGCC 1500
Db 1494 GATTTCTCAACAGCGTGCATTCGTGGAAGACATTAATAGGGCTTGTGATTTGGCC 1553
Qy 1501 CCAGAGATATACATGATTTCTCGAGTACAGTGA 1533
Db 1554 CCAGAGATATACATGATTTCTCGAGTACAGTGA 1586
```

```
RESULT 11
AAV62440
ID AAV62440 standard; cDNA; 1782 BP.
```

```
XX AAV62440;
```

```
DT 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
```

```
XX Soybean wild-type myo-inositol 1-phosphate synthase cDNA.
```

```
KW Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KW phytic acid; ds.
```

```
OS Glycine max; line LR13.
```

```
PH Key Location/Qualifiers
FT CDS 54..1586
/*tag= a
```

```
XX MO9845448-A1.
```

```
XX 15-OCT-1998.
```

```
XX 07-APR-1998; 98WO-US006822.
```

```
XX 08-APR-1997; 97US-00835751.
```

```
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
```

```
XX Hitz WD, Sebastian SA;
```

```
XX WPI; 1998-568353/48.
```

```
XX P-PSDB; AAW79740.
```

```
XX Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
```

```
PT phytic acid, etc.
```

```
XX Example 5; Page 44-45; 63pp; English.
```

```
XX This is the nucleotide sequence of cDNA encoding the wild-type soybean
PS myo-inositol 1-phosphate synthase (MI 1-PS) present in clone p5sm1-1ps
XX (ATCC 97970). The clone was isolated from a cDNA library of soybean line
CC LR13 by hybridisation to a probe made from MI 1-PS cDNA of Arabidopsis
CC thaliana. MI 1-PS is involved in glucose metabolism to phytic acid,
CC raffinose and stachyose. A mutant MI 1-PS nucleic acid (see AAV62443) has
CC been identified in soybean line LR13, a mutagenised line of low raffinose
CC saccharide phenotype. Sequencing revealed a single base change mutation
CC (G to T at base 1241) in the LR13 sequence. The mutation results in a
CC seed phenotype of very low raffinose saccharide sugars, very high sucrose
CC and low phytic acid. The nucleic acid is used to alter the raffinose
CC saccharide, sucrose, phytic acid and inorganic phosphate content of
CC soybean seeds, leading to useful soybean products, e.g. a seed phytic
CC acid content of less than 17 ug/g, a seed content of raffinose and
CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
CC greater than 200 ug/g. (updated on 17-OCT-2003 to standardise OS field)
```

```
XX Sequence 1782 BP; 516 A; 371 C; 411 G; 484 T; 0 U; 0 Other;
```

```
Qy Query Match 95.5%; Score 1464.2; DB 2; Length 1782;
```

```
Db Best Local Similarity 97.2%; Pred. No. 0;
```

```
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
```

```
Qy 1 ATGTTCAATCGAAGATTTTAAGGTAGAGTCTTAATGTAAGTACACCGAGCTGAGATT 60
Db 54 ATGTTCAATCGAAGATTTTAAGGTAGAGTCTTAATGTAAGTACACCGAGCTGAGATT 113
Qy 61 CAGTCCGTTGTAACACTAGAAACCAACGAACTTTTCAAGAAACAGAAATGACCTTAT 120
Db 114 CAGTCCGTTGTAACACTAGAAACCAACGAACTTTTCAAGAAACAGAAATGACCTTAT 173
Qy 121 CAGTGGATTGTCAAAACCAATCCGTCAACTACCAATTTAAACCAACCAATTTTCCA 180
Db 174 CAGTGGATTGTCAAAACCAATCCGTCAACTACCAATTTAAACCAACCAATTTTCTT 233
Qy 181 AAATTGGGGGTGATGCTTTGGGGTTGGGGTGAACCAACGCTTACCTCAACCGGTGAT 240
Db 234 AAATTAGGGGTATATCTTTGGGGTGGGGTGAACCAACGCTTACCTCAACCGGTGAT 293
Qy 241 GTTATTTGCTAACAGAGAGACATTTTATGAGGTACAAAGAGCAAGATTTCAACCAAT 300
Db 294 GTTATTTGCTAACAGAGAGACATTTTATGAGGTACAAAGAGCAAGATTTCAACCAAT 353
Qy 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCAGATTGGATCTTCCAGGGAAGAGA 360
Db 354 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTCAGATTGGATCTTCCAGGGAAGAGA 413
Qy 361 ATCTATGCCCATTCAGAGATGCTGCTTCAATGTTAAATCTGACGACATTTGTTGGG 420
Db 414 ATCTATGCCCATTCAGAGATGCTGCTTCAATGTTAAATCTGACGACATTTGTTGGG 473
Qy 421 GGATGGGATATCAGCAATGAACTGGGTATGCTATGCTGATGCGCAAGGCAAAAGTGTGAC 480
Db 474 GGATGGGATATCAGCAATGAACTGGGTATGCTATGCTGATGCGCAAGGCAAAAGTGTGAC 533
Qy 481 ATGATTTTGCAGAGCAGTTGAGGCTTTTACATGAAATCATGTTCCACTCCCGGAATC 540
Db 534 ATGATTTTGCAGAGCAGTTGAGGCTTTTACATGAAATCATGTTCCACTCCCGGAATC 593
Qy 541 TACGACCCGGATTTTATGCTGCTCAACCAAGAGGAGCGTCCCAACAGTGTAAAGGCG 600
Db 594 TATGACCCGGATTTTATGCTGCTCAACCAAGAGGAGCGTCCCAACAGTGTAAAGGCG 653
Qy 601 ACAAGCAAGAGCAAGTTCAGCAATCATCAAGACATCAAGGCTTTTAAGGAACCAAC 660
Db 654 ACAAGCAAGAGCAAGTTCAGCAATCATCAAGACATCAAGGCTTTTAAGGAACCAAC 713
Qy 661 AAATGGAACAAAGTGTGCTGTTGACTGCAACACAGAGAGGTATGCAATTTGGTT 720
```

[illegible]

OS	Glycine max; line LR33.
XX	
PN	W09845448-A1.
PD	
XX	15-OCT-1998.
XX	
PF	07-APR-1998; 98MO-US006822.
PR	
XX	08-APR-1997; 97US-00835751.
PA	(DUPO) DU PONT DE NEMOURS & CO E I.
XX	
PI	Hitz WD, Sebastian SA;
DR	
DR	WPI, 1998-568353/48.
XX	P-PsDB; AAW79741.
PT	
PT	Soybean plants containing altered myo-inositol-1-phosphate gene - useful for generating plants wth altered levels of e.g. raffinose, stachyose, phytic acid, etc.
XX	
PS	Example 5; Page 48-49; 63pp; English.
XX	
CC	This is the nucleotide sequence of cDNA encoding a mutant soybean myo-
CC	inositol 1-phosphate synthase (MI 1-PS) present in clone LR33-10. This
CC	clone was isolated from a CDNA library of soybean line LR33 by PCR
CC	amplification (see AAV62441-42). Line LR33 was obtained by chemical
CC	mutagenesis of wild-type soybean genome and as a reduced raffinose
CC	saccharide phenotype. MI 1-PS is involved in glucose metabolism to phytic
CC	acid, raffinose and stachyose. Sequencing revealed a single base change
CC	muation (G to T at base 1241) in the LR33 sequence when compared to the
CC	wild-type sequence (see AAV62440). The mutation results in a seed
CC	phenotype of very low raffinose saccharide sugars, very high sucrose and
CC	low phytic acid. The mutated MI 1-PS nucleic acid is used to alter the
CC	raffinose saccharide, sucrose, phytic acid and inorganic phosphate
CC	content of soybean seeds, leading to useful soybean products, e.g. a seed
CC	phytic acid content of less than 17 ug/g, a seed content of raffinose and
CC	stachyose combined of less than 14.5 ug/g, and a seed sucrose content
CC	greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)
XX	
SO	Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;
	Query Match 95.4%; Score 1462.6; DB 2; Length 1533;
	Best Local Similarity 97.1%; Pred. No. 0;
	Matches 1489; Conservative 0; Mismatches 44; Indels 0; Gaps 0.
OY	1 ATGTTCATCGAANTTTTAAGTGAAGAAGTCTTAATGTGAAGTACCCGAGACTGAGATT 60
DB	1 ATGTTCATCGAANTTTTAAGTGAAGTCTTAATGTGAAGTACCCGAGACTGAGATT 60
OY	61 CAGTCCGGTGCACTGTCGAAACCAACCGAATTCTTACAAGGAAACAGGAATGGACCATT 120
DB	61 CAGTCCGGTGCACTGTCGAAACCAACCGAATTCTTACAAGGAAACAGGAATGGACCATT 120
OY	121 CAGTGAATTGCACAACCCAATCCGTCACATCACCATTAAAACCAACCCATATGTTTCCA 180
DB	121 CAGTGAATTGCACAACCCAATCCGTCACATCACCATTAAAACCAACCATATGTTTCTT 180
OY	181 AAATTGGGGGTGATGCTTGTGGGTTGTGGGTTGGAACAACGCGCTCTACCTCACCGGTGT 240
DB	181 AAATTGGGGGTGATGCTTGTGGGTTGTGGGTTGGAACAACGCGCTCTACCTCACCGGTGT 240
OY	241 GTTATGTCGTAACAAGAGGACATTTCAATGGGCTTCAAAGGACAGAATTCACAAGGCCAAT 300
DB	241 GTTATGTCGTAACAAGAGGACATTTCAATGGGCTTCAAAGGACAGAATTCACAAGGCCAAT 300
OY	301 TACTTTGGCTCCCTCACCCCAAGCCTCAGACTATTGAGTTGGATCTCTTCAGAGGAGAGGAA 360
DB	301 TACTTTGGCTCCCTCACCCCAAGCCTCAGACTATTGAGTTGGATCTCTTCAGAGGAGAGGAA 360
OY	361 ATTATATGCCCATTCANAAGAGTGTGCTTCAATGGTAAATCTTGACGACATTTGTGTGGG 420
DB	361 ATTATATGCCCATTCANAAGAGTGTGCTTCAATGGTAAATCTTGACGACATTTGTGTGGG 420

QY 421 GGATGGATATATGAGCAACATGAACCTGGCTGATGCCATGCGCAGGSCAAAGGTGTTTGC 480
 DB 421 GGATGGATATATGAGCAACATGAACCTGGCTGATGCCATGCGCAGGSCAAAGGTGTTTGC 480
 QY 481 ATGATTTGACAGAGAGAGTTGAGGCTTATCATGAAATCCATGCTTCCCTCCGGAATC 540
 DB 481 ATGATTTGACAGAGAGAGTTGAGGCTTATCATGAAATCCATGCTTCCCTCCGGAATC 540
 QY 541 TACGACCCGGATTTATGCTGCTCCAAACGAGGAGCGGCCAACAAAGTGAATTAAGAGGC 600
 DB 541 TATGACCCGGATTTATGCTGCTCCAAACGAGGAGCGGCCAACAAAGTGAATTAAGAGGC 600
 QY 601 ACAAGCAAGACGCAAGTTTCAGCAAAATCATCAAAAGCATCAAGCGCTTTAAGAGACCAAC 660
 DB 601 ACAAGCAAGACGCAAGTTTCAGCAAAATCATCAAAAGCATCAAGCGCTTTAAGAGACCAAC 660
 QY 661 AAAGTGACAGAGGTGTTCTCTGAGACTGCGCAACACAGAGAGTATGCAATTTGGTT 720
 DB 661 AAAGTGACAGAGGTGTTCTCTGAGACTGCGCAACACAGAGAGTATGCAATTTGGTT 720
 QY 721 GTAGGCTTATGACACCATGAGAAATCTTGGCTGCTGAGAGAGAAATGAGGCTGAG 780
 DB 721 GTAGGCTTATGACACCATGAGAAATCTTGGCTGCTGAGAGAGAAATGAGGCTGAG 780
 QY 781 ATTTCTCTTCCACCTTGATGCAATGCTGCTGATGAGAAATGTTCTTTCAATTAAT 840
 DB 781 ATTTCTCTTCCACCTTGATGCAATGCTGCTGATGAGAAATGTTCTTTCAATTAAT 840
 QY 841 GGAAGCCCTCAGAAACATTTTGTACAGAGGCTGATGATCTTCCATCGCGAGAACACT 900
 DB 841 GGAAGCCCTCAGAAACATTTTGTACAGAGGCTGATGATCTTCCATCGCGAGAACACT 900
 QY 901 TTGATTTGGTGAAGTACTTCAAGAGTGGTCAGACCAAAATGAATCTGTTGGTGAAT 960
 DB 901 TTGATTTGGTGAAGTACTTCAAGAGTGGTCAGACCAAAATGAATCTGTTGGTGAAT 960
 QY 961 TTTCTTGGGGGCTGGTATCAAGCCCAATCTATAGTTAGTAAACAACCATCTGGGAAAC 1020
 DB 961 TTTCTTGGGGGCTGGTATCAAGCCCAATCTATAGTTAGTAAACAACCATCTGGGAAAC 1020
 QY 1021 AATGATGATATATCTCTGGCTCCAAACCTTCCGCTCCAGAGAAATCTTCCAAAGC 1080
 DB 1021 AATGATGATATATCTCTGGCTCCAAACCTTCCGCTCCAGAGAAATCTTCCAAAGC 1080
 QY 1081 AAGCTTTGAGATATGATGCAACAGCAATGCCATCTCTATAGAGCTGGTGAACATCCA 1140
 DB 1081 AAGCTTTGAGATATGATGCAACAGCAATGCCATCTCTATAGAGCTGGTGAACATCCA 1140
 QY 1141 GACCATGTTGTTATTAAGTATGCTCTTACGTAGGAGATAGCAAGAGCCCATGAT 1200
 DB 1141 GACCATGTTGTTATTAAGTATGCTCTTACGTAGGAGATAGCAAGAGCCCATGAT 1200
 QY 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAAACACATTTGTTTGCACACATAT 1260
 DB 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAAACACATTTGTTTGCACACATAT 1260
 QY 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAAACACATTTGTTTGCACACATAT 1260
 DB 1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAAACACATTTGTTTGCACACATAT 1260
 QY 1261 GAGGATTCCTTTTACCTGCTCTATTAATCTTGAATGATGCTTCTTCTTCTGAGCTCAGC 1320
 DB 1261 GAGGATTCCTTTTACCTGCTCTATTAATCTTGAATGATGCTTCTTCTTCTGAGCTCAGC 1320
 QY 1321 ACTAGATTCAGTTTAAAGCTGAAAATGAGGAAAATTCACATCTTCCACCCAGTTGCT 1380
 DB 1321 ACTAGATTCAGTTTAAAGCTGAAAATGAGGAAAATTCACATCTTCCACCCAGTTGCT 1380
 QY 1381 ACCATTTCTAGCTATATGACCAAGGCTCTCTGTTTCCACCGGTTACACCAAGTGGTGAAT 1440
 DB 1381 ACCATTTCTAGCTATATGACCAAGGCTCTCTGTTTCCACCGGTTACACCAAGTGGTGAAT 1440
 QY 1441 GGAATGTCAAAGAGGCTGCATGCTGGAACCATTAATAGGGCTGTTGTTGGAATTTGGCC 1500
 DB 1441 GGAATGTCAAAGAGGCTGCATGCTGGAACCATTAATAGGGCTGTTGTTGGAATTTGGCC 1500

QY 1501 CCAGAGAAATAACATGATTTCTCGAGTACAGTGA 1533
 DB 1501 CCAGAGAAATAACATGATTTCTCGAGTACAGTGA 1533
 RESULT 13
 ID AD014494
 AC AD014494 standard; cDNA; 1533 BP.
 XX
 AC AD014494;
 XX
 DT 23-SEP-2004 (first entry)
 XX
 DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #1.
 XX
 KW Soybean; myo-inositol 1-phosphate synthase; gene; *ss*;
 KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
 KM inorganic phosphate; mutant.
 XX
 OS Glycine max.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1533
 FT /*tag= a
 FT /product= "Mutant soybean myo-inositol 1-phosphate
 FT mutation synthase #1"
 FT replace(1241,G)
 FT /*tag= b
 PN US2004128713-A1.
 PD 01-JUL-2004.
 XX
 PF 21-NOV-2003; 2003US-00718952.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 PR 26-APR-1999; 99US-0029315.
 PR 11-MAR-2002; 2002US-00025003.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PA (GRAC/) GRACE D J.
 PA (STRE/) STREIT L G.
 XX
 PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
 XX
 DR WPI; 2004-533135/51.
 DR P-PSDB; AD014495.
 XX
 PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
 PT useful for producing plants with decreased raffinose, stachyose, and
 PT phytic acid and increased sucrose, leading to valuable and useful soybean
 PT products.
 XX
 PS Claim 10; SEQ ID NO 5; 48bp; English.
 XX
 CC The invention relates to a nucleic acid fragment encoding a soybean myo-
 CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
 CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
 CC phosphate. The invention also relates to a chimeric gene operably linked
 CC to suitable regulatory sequences, where expression of the chimeric gene
 CC results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
 CC comprising the chimeric gene, a method of making the soybean plant, a
 CC seed of the soybean plant, a soy protein product derived from the
 CC processing of soybean seeds, a method of making or producing a soy
 CC protein product and a method of using a soybean plant homozygous for at
 CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
 CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
 CC nucleic acid fragment and methods are useful for producing plants with
 CC decreased raffinose, stachyose and phytic acid content and increased

CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1533 BP, 429 A, 344 C, 367 G, 393 T, 0 U, 0 Other;

Query Match 95.4%; Score 1462.6; DB 12; Length 1533;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1489; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```

QY 1 ATGTCATCGAGAAATTTAAAGTAGAGAGTCTTAATGTGAATGACCGAGACTGAGATT 60
DB 1 ATGTCATCGAGAAATTTAAAGTAGAGTCTTAATGTGAATGACCGAGACTGAGATT 60
QY 61 CAGTCCGTGTACAACTACCAACCACTGTTTCAAGAAACAGAAATGGCACTTAT 120
DB 61 CAGTCCGTGTACAACTACCAACCACTGTTTCAAGAAACAGAAATGGCACTTAT 120
QY 121 CAGTGGATTGTCAAAACCAATCCGTCACTACCAATTTAAACCAACCACTGTTTCA 180
DB 121 CAGTGGATTGTCAAAACCAATCCGTCACTACCAATTTAAACCAACCACTGTTTCT 180
QY 181 AAATTGGGGGTGATGCTTGTGGGTGGGTGAAACACGCTCTACCTCACCCTGAGT 240
DB 181 AAATTGGGGGTATGCTTGTGGGTGGGTGAAACACGCTCACCCTCACCCTGAGT 240
QY 241 GTTATTGCTTACAGAGACATTTTATGAGGCTTACAAAGACAAAGATTCAACGCCAAT 300
DB 241 GTTATTGCTTACAGAGACATTTTATGAGGCTTACAAAGACAAAGATTCAACGCCAAT 300
QY 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGAGTCTTCCAGGGAGAGAA 360
DB 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTGAGTTGAGTCTTCCAGGGAGAGAA 360
QY 361 ATCTATGCCCCCATTTCAAGAGTCTGCTTCAATGTTTATCTTGCACATTTGTTGGG 420
DB 361 ATCTATGCCCCCATTTCAAGAGCTGCTTCCAAATGTTTATCCTTGCACATTTGTTGGG 420
QY 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCCATGGCCAAAGGTTTGTAC 480
DB 421 GGATGGGATATCAGCAACATGAACTGGCTGATGCCATGGCCAAAGGTTTGTAC 480
QY 481 ATCGATTTGTCAGAAAGAGTTGAGGCTTTCATGGAATTCATGCTCCCGGGAATC 540
DB 481 ATCGATTTGTCAGAAAGAGTTGAGGCTTTCATGGAATTCATGCTCCCGGGAATC 540
QY 541 TACGACCCCGGATTTTCTGCTGCAACCAAGAGACGTCGCAACACGTTAAGAGGC 600
DB 541 TATGACCCCGGATTTTCTGCTGCAACCAAGAGACGTCGCAACACGTTAAGAGGC 600
QY 601 ACAAGACAAAGACAGTTCAACAAATCATCAAGATCAAGCGCTTTAAGAGACACC 660
DB 601 ACAAGACAAAGACAGTTCAACAAATCATCAAGATCAAGCGCTTTAAGAGACACC 660
QY 661 AAAAGTGACAAGGTGGTGTCCGTTGAGTGCACCAAGAGAGTATGCAATTTGTT 720
DB 661 AAAAGTGACAAGGTGGTGTCTGTTGAGTGCACCAAGAGAGTATGCAATTTGTT 720
QY 721 GTAGGCTTAAATGACACATGAGAAATCTTTGCTGCTGTGGACAGAAATGAGCTGAG 780
DB 721 GTAGGCTTAAATGACACATGAGAAATCTTTGCTGCTGTGGACAGAAATGAGCTGAG 780
QY 781 ATTTCTCTTCCACTTGTATGCAATGCTGCTGTGTGATGAAAATGTTCTTTCAATAT 840
DB 781 ATTTCTCTTCCACTTGTATGCAATGCTGCTGTGTGATGAAAATGTTCTTTCAATAT 840
QY 841 GGAAGCCCTCAGAAACCTTTGACACAGGGCTGATGATCTTGCCATGCGCGAGAACT 900
DB 841 GGAAGCCCTCAGAAACCTTTGACACAGGGCTGATGATCTTGCCATGCGCGAGAACT 900
QY 901 TTGATTGTGTGAGATGACTTCAAGAGTGTGACCAACCAATGAATCTGTGTTGTTGAT 960
DB 901 TTGATTGTGTGAGATGACTTCAAGAGTGTGACCAACCAATGAATCTGTGTTGTTGAT 960

```

```

QY 961 TTCTCTTGGGGGCTGGTATCAAGCCACATCTATGATTAGTACCAACATCTGGGAAC 1020
DB 961 TTCTCTTGGGGGCTGGTATCAAGCCACATCTATGATTAGTACCAACATCTGGGAAC 1020
QY 1021 AATGATGTATGATATCTCTGCGCTCCACAAACCTTCCGCTCCAGAGAAATCTTCAAGAC 1080
DB 1021 AATGATGTATGATATCTTCCGCTCCACAAACCTTCCGCTCCAGAGAAATCTTCAAGAC 1080
QY 1081 AACGTTGTTGACGATATGTTCAACAGCAATGCCATCTCTATGAGCTGTGGTAACTCC 1140
DB 1081 AACGTTGTTGATATGTTGTTCAACAGCAATGCCATCTCTATGAGCTGTGGTAACTCC 1140
QY 1141 GACATGTTGTTGTTATTAAGTATGCTTACGTACGTAGGAGATGACAAAGAGCCATGAT 1200
DB 1141 GACCATGTTGTTGTTATTAAGTATGCTTACGTACGTAGGAGATGACAAATGAGCCATGAT 1200
QY 1201 GAGTACACTTCAGAGATATTCATGAGGTGGAAGAAACACCATTTGTTGCAACACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTCATGAGGTGGAAGAAACACCATTTGTTGCAACACATGT 1260
QY 1261 GAGGATTCCTTTTATGCTGCTCTATTTATCTTGAATTTGACTTCTTCTGAGCTGAC 1320
DB 1261 GAGGATTCCTCTTATGCTGCTCTATTTATCTTGAATTTGACTTCTTCTGAGCTGAC 1320
QY 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCATCTATTTCCACCAAGTTGCT 1380
DB 1321 ACTAGATTCAGTTTAAAGCTGAAATGAGGAAATTCATCTATTTCCACCAAGTTGCT 1380
QY 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTGTTTCCACCGGGTACACAGTGGTAAAT 1440
DB 1381 ACCATTTCTAGCTATCTGACCAAGGCTCTGTTTCCACCGGGTACACAGTGGTAAAT 1440
QY 1441 GCATTGTCAAGACAGTGCATATGCTGAAACCATTAATGAGGCTTGTGATTTGGCC 1500
DB 1441 GCATTGTCAAGACAGTGCATATGCTGAAACCATTAATGAGGCTTGTGATTTGGCC 1500
QY 1501 CCAGAGATTAATGATTTCTGAGTACAAAGTGA 1533
DB 1501 CCAGAGATTAATGATTTCTGAGTACAAAGTGA 1533

```

RESULT 14

ADS81997 standard; cDNA, 1533 BP.

ADS81997;

18-NOV-2004 (first entry)

Soybean myo-inositol 1-phosphate synthase cDNA mutant #1.

Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.

Glycine max; line LR33.

OS Synthetic.

Key Location/Qualifiers

FT 1..1533

FT /*tag= a

FT /product= "myo-inositol 1-phosphate synthase"

FT replace(1188,G)

FT /*tag= b

PN US2003074685-A1.

PD 17-APR-2003.

PF 11-MAR-2002; 2002US-00025003.

PR 08-APR-1997; 97US-00835751.

PR 07-APR-1996; 98WO-US006822.

DB 1381 ACCATCTCAGTACTACCAAGGCTCTCTGTTCCACCGGGTACACAGTGTGAAT 1440
QY 1441 GCATTGTCGAAGAGGCTGCATGTCGAAACATATATGAGGGCTGTGTGGATTGGCC 1500
DB 1441 GCATTGTCGAAGAGGCTGCATGTCGAAACATATATGAGGGCTGTGTGGATTGGCC 1500
QY 1501 CCAGAGATTAACATGATTCTCGATCAAGTGA 1533
DB 1501 CCAGAGATTAACATGATTCTCGATCAAGTGA 1533
RESULT 15
AAK90402
ID AAK90402 standard; cDNA to mRNA, 1950 BP.
XX AAK90402;
XX
XX 24-SEP-1999 (first entry)
XX
XX Nicotiana paniculata INPS encoding cDNA.
XX
XX Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;
XX water stress; resistance; ds.
XX Nicotiana paniculata.
XX JP1187879-A.
XX
XX 13-JUL-1999.
XX
XX 26-DEC-1997; 97JP-00359773.
XX
XX 26-DEC-1997; 97JP-00359773.
XX
XX (NISR) JAPAN TOBACCO INC.
XX
XX WPI; 1999-451546/38.
XX P-PSDB; AAY24477.
XX
XX New INPS gene derived from Nicotiana genus plant - useful for conferring
XX resistance to water stress to plants.
XX
XX
XX Claim 2; Page 6-8; Bpp; Japanese.
XX
XX The present sequence encodes Nicotiana paniculata inositol monophosphate
XX synthase (INPS), designated NpINPS1. INPS can be used to confer water
XX stress resistance to a plant
XX
XX Sequence 1950 BP; 570 A; 377 C; 425 G; 578 T; 0 U; 0 Other:
SQ
Query Match 71.6%; Score 1097.8; DB 2; Length 1950;
Best Local Similarity 82.3%; Pred. No. 0;
Matches 1261; Conservative 0; Mismatches 272; Indels 0; Gaps 0;
QY 1 AGGTCATGAGAAATTTAAGGTAGAGAGTCCATAGTGAAGTACCGAGCTGAGATT 60
DB 92 AGGTTATTAGAGAAATTTAAGGTAGAGAGTCCATAGTGAAGTACCGAGTGAAT 151
QY 61 CAGTCGTGTACAACTACGAAACCAACGAACTGTTCAGAGAAACAGATGSCACTAT 120
DB 152 CACTCGTCTATGATTATCAACCACTGATGTTATCATGATGAGAAATGSGACATAT 211
QY 121 CAGTGTATGTCAAAACCAATCCGTCATCTACCAATTTAAACCAACCACTGTTCA 180
DB 212 CAGTGAACGTCAGAGCTAAGACTGTCAATATGAGTTCAAGCTGATGTTATCTCC 271
QY 181 AATTTGGGGGTATGCTTGTGGGTGGGTAACAAGGCTCTACCCCTACCGGTGAT 240
DB 272 AATTTGGGGGTATGCTTGTGGGTGGGTAACAAGGCTCTACCCCTACCGGTGAT 331
QY 241 GTTATTTCTAACAGAGACATTTTCATGAGGCTTACAAAGACAAAGATTCAAGCCAT 300
DB 332 GTTATTTCTAACAGAGACATTTTCATGAGGCTTACAAAGACAAAGATTCAAGCCAT 391

QY 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTTCGAGTTGATTCCTTCCAGGAGAGAA 360
DB 392 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTTCGAGTTGATTCCTTCCAGGAGAGAA 451
QY 361 ATCTATGCCCAATTCGAAGAGTCTGCTTCCATATGTTATCTGACGACATGTTGTTGG 420
DB 452 ATCTATGCTCCATTTTAAAGCTCTCTTCCATATGTTATCTGACGATACGATGTTTGA 511
QY 421 GGAATGGATATCAGCAACATGAACCTGCTGATGCTGATGCTGCAAGGCAAGGTTTGA 480
DB 512 GGAATGGATATCAGCAACATGAACCTGCTGATGCTGATGCTGCAAGGCAAGGTTTGA 571
QY 481 ATGATTTGACAGAGAGATGAGGCTTTCATGAAATTCATGTTTCACTCCCGCAATC 540
DB 572 ATGATTTGACAGAGAGATGAGGCTTTCATGAAATTCATGTTTCACTCCCGCAATC 631
QY 541 TAGACCCGGATTTTCATGCTGCTGCAACCAAGAGAGGCTGCAACCAAGTATTAAGG 600
DB 632 TAGACCCGGATTTTCATGCTGCTGCAACCAAGAGAGGCTGCAACCAAGTATTAAGG 691
QY 601 ACAAGCAAGACCAAGTTCAGCAAAATCATCAAGACATGAGGCTTTAAGAGACACC 660
DB 692 ACAAGCAAGACCAAGTTCAGCAAAATCATCAAGACATGAGGCTTTAAGAGAGAAC 751
QY 661 AAGTGAACAGAGTGTGCTCTGCTGACCTGCAACAGAGAGATATGCAATTTGTT 720
DB 752 AAGTGAACAGAGTGTGCTCTGCTGACCTGCAACAGAGAGATATGCAATTTGTT 811
QY 721 GTAGGCTTAATGACCAATGAGAGATCTTGGCTGCTGAGCAAGAAATGAGGCTGAG 780
DB 812 GTTGAATTAATGACCAATGAGAGATCTTGGCTGCTGAGCAAGAAATGAGGCTGAG 871
QY 781 ATTTCTCTTCCACCTTGATGCAATGCTGCTGATGAGAAATGTTCTTTCATTAAT 840
DB 872 ATATCTCTTCCACCTTGATGCAATGCTGCTGATGAGAAATGTTCTTTCATTAAT 931
QY 841 GGAAGCCCTCAACCACTTTTATCAAGGCTGATGATCTTCCATCCGAGAACACT 900
DB 932 GGAAGCCCTCAACCACTTTTATCAAGGCTGATGATCTTCCATCCGAGAACACT 991
QY 901 TTGATTTGATGAGATGATCTTCAAGAGTGTGCAACCAAAATGTTGTTGTTGAT 960
DB 992 TTGATTTGATGAGATGATCTTCAAGAGTGTGCAACCAAAATGTTGTTGTTGAT 1051
QY 961 TTTCTTGGGGCTGTATCAAGCCCAATCTATATGTTAGTCAACCATCTGGGAAAC 1020
DB 1052 TTTCTTGGGGCTGTATCAAGCCCAATCTATATGTTAGTCAACCATCTGGGAAAC 1111
QY 1021 AATGATGATGATATCTCTGCTTCAACCACTTCCGCTCAAGAAATTTCCAGAGC 1080
DB 1112 AATGATGATGATATCTCTGCTTCAACCACTTCCGCTCAAGAAATTTCCAGAGC 1171
QY 1081 AAGCTTTGAGATATGATGATCAACGCAATGCAATCTTATGAGCTGAGTGAACATCC 1140
DB 1172 AATGTTGATGATATGATGATCAACGCAATGCAATCTTATGAGCTGAGTGAACATCC 1231
QY 1141 GACATGTTGTTGTTATTAAGTATGCTTCAAGAGGATGAGAGAGACCATGAT 1200
DB 1232 GACATGTTGTTGTTATTAAGTATGCTTCAAGAGGATGAGAGAGACCATGAT 1291
QY 1201 GAGTACATTCAGAGATATTCATGAGTGAAGAAACACCATTTGTTGCAACACATGT 1260
DB 1292 GAGTACATTCAGAGATATTCATGAGTGAAGAAACACCATTTGTTGCAACATGT 1351
QY 1261 GAGATTCCTTTGATGCTGCTTATTAATCTTGAATGCTTCTTCTGAGCTGAGC 1320
DB 1352 GAGATTCCTTTGATGCTGCTTATTAATCTTGAATGCTTCTTCTGAGCTGAGC 1411
QY 1321 ACTGAATTCAGTTTAAAGCTGAAATGAGGAAATTCATCTATTCACCCAGTTGCT 1380
DB 1412 ACCGCATTCAGCTCAAGGCTGAAGAGAGGTTAAGTTCACTCTTCAACCCGTTGCT 1471

Qy 1381 ACCATTCTGAGCTATCTGACCAAGGCTCCTTGTTCCACCGGGTACACCACTGTTGAAT 1440
Db 1472 ACTATCCTCAGCTACCTTACCAAGGCTCCTTGTTACACCAAGTACACCAAGTGAAT 1531
Qy 1441 GCATTGTCAAAGCAGGCTGCAATGCTGAAAACATATATGAGGCTTGTGTGATTGGCC 1500
Db 1532 GCACTCTCAAAGCAGGCGCAATGCTTGAGAACATATATGAGGCTTGTGTGACTTGCA 1591
Qy 1501 CCAGAGATATACATGATTTCTCGAGTACAAGTGA 1533
Db 1592 CCAGAGAACACATGATTTCTGGAATACAATGA 1624

Search completed: June 7, 2005, 22:54:28
Job time : 850.546 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 20:51:42 ; Search time 6686.92 Seconds
(without alignments)
11108.537 Million cell updates/sec

Title: US-10-718-952-11

Perfect score: 1533

Sequence: 1 atgttcacgcgaattttaa.....tgattctcgagtaacagfga 1533

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmb1:*
1: gb_ha:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1464.2	95.5	1739	8	AY038802 Glycine m
2	1464.2	95.5	1782	8	BD075266 Soybean p
3	1462.6	95.4	1533	6	BD075269 Soybean p
4	1456.2	95.0	1791	8	AF293970 Glycine m
5	1097.8	71.6	1931	8	AB032073 Nicotiana
6	1097.8	71.6	1950	6	E27176 Novel INPS
7	1093	71.3	1959	8	AB009881 Nicotiana
8	1089	69.7	1954	8	BT013759 Lycopersi
9	1061	68.2	1959	8	BT013505 Lycopersi
10	1045	68.2	1845	8	AF284065 Sesamum i
11	1044.6	68.1	1978	8	CPINOLG
12	1027.4	67.0	2053	8	MCHJ2511
13	1020.4	66.6	1986	8	AF433879 Suaeda ma
14	1016.2	66.3	1845	8	AY028259 Avicennia
15	1000.2	65.2	1781	6	AX054630 Sequence
16	1000.2	65.1	1781	6	BN06307 Braessia na
17	998.6	65.1	1533	6	CO805008 Sequence
18	998.6	65.1	1533	6	AX506743 Sequence
19	998.6	65.1	1533	8	AY143904 Arabidops

20	998.6	65.1	1863	8	AY054202 Arabidops
21	998.6	65.1	1864	8	AY053415 Arabidops
22	997	65.0	1538	8	AF120146 Triticum
23	997	65.0	1538	8	AF120147 Triticum
24	997	65.0	1538	8	AF120148 Triticum
25	990.6	64.6	1564	8	AY096554 Arabidops
26	990.6	64.6	1564	8	AY096554 Arabidops
27	989	64.5	1890	8	ATU30250 Arabidops
28	987.2	64.4	1870	8	AF372954 Arabidops
29	985.8	64.3	1535	6	BD073472 Regulated
30	985.8	64.3	1535	6	BD073472 Regulated
31	985.8	64.3	1665	8	AF056326 Zea mays
32	985.6	64.3	1567	8	BT001931 Arabidops
33	985.6	64.3	1837	8	AY085931 Arabidops
34	984.2	64.2	1533	8	AY323824 Xerophyta
35	984.2	64.2	1959	6	AX151207 Sequence
36	982.4	64.1	1921	8	ATU04876 Arabidops
37	981	64.0	1914	8	AK103501 Oryza sat
38	981	64.0	1915	8	AK058750 Oryza sat
39	979.4	63.9	1868	8	AB012107 Oryza sat
40	979.4	63.9	1931	6	AR137808 Sequence
41	979.4	63.9	1931	6	AR170167 Sequence
42	976	63.7	1871	8	PVU38920 Phaseolus v
43	974.6	63.6	1936	8	AB059557 Avena sat
44	973.2	63.5	1930	8	AY035045 Arabidops
45	969.8	63.3	2152	8	AF056325 Hordeum v

ALIGNMENTS

RESULT 1
LOCUS AY038802 1739 bp mRNA linear PLN 16-JUL-2001
DEFINITION Glycine max myo-inositol-1-phosphate synthase mRNA, complete cds.
ACCESSION AY038802
VERSION AY038802.1 GI:14764465

KEYWORDS
SOURCE
ORGANISM
Glycine max (soybean)

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE
1 (bases 1 to 1739)
Hitz,W.D., Carlson,T.J., Kerr,P. and Sebastian,S.
Biochemical and Molecular Characterization of a Mutation that Confers a Decreased Raffinose and Phytic Acid Phenotype on Soybean Seeds

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Unpublished
2 (bases 1 to 1739)
Carlson,T.J. and Hitz,W.D.
Direct Submission
Submitted (06-JUN-2001) Crop Genetics, DuPont Co., P.O.Box 80402, Wilmington, DE 19880-0402, USA

FEATURES
source
Location/Qualifiers
1..1739

/organism="Glycine max"
/mol_type="mRNA"
/culтивар="Wye"
/db_xref="taxon:3847"
/dev_stage="developing embryo"
11..1543
/EC_number="5.5.1.4"
/function="cyclizes glucose 6-phosphate to l-myo-inositol-1-phosphate"
/note="m11PS-1A"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAK72098.1"
/db_xref="GI:14764466"
/translation="MTFENPVEGCVNYTETIYSVYVETTELVHNRNGTYOMIV
KPKSVYEFKTNHVPKLGVMVGNGNGSTLTGVIANREGISMATKDKIQQANYF

GSUFOASAIRVGSFOGEEIYAPFKSLPLPVNPDIVFGMDISNNMLADAMARAFVD
IDLKOLRPMESMLPLGIYDPDRIANNOBERANNVJKTROBOVOOIIDIXAFKE
ATKXRVVVLMTAMTERYSNLVYGNDRNENLLAVDRNEAISTSTYAIACVENV
PFINSPOFTVEPLIDILAIANTLIGDDEFSGOTKMSVLDVLVAGAKPISIVS
YNHLKNDGOMNLSAPOTFRSKISKSNVDVNSNALIYEGEHPDHVAVKPYV
GDSKAMBEYSEIEMGKSTIVLHNTCEDSLAAILDVLIALESTRLEFKENE
KGFHSPHVAITLISYLTKAPLIVPGPIVNVNLSKORAMLENIMRACVGLAPENNMLE
YK"

ORIGIN

Query Match 95.5%; Score 1464.2; DB 8; Length 1739;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 ATGTTTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAATACCGAGACTGAGATT 60
DB 11 AGTTTCATCGAAGAAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATT 70
QY 61 CAGTCGGGTAGACATGAGAAACACCGAACTTGTTCAGAGAACGAGATGCACTAT 120
DB 71 CAGTCGGGTAGACATGAGAAACACCGAACTTGTTCAGAGAACGAGATGCACTAT 130
QY 121 CAGTGGATGTCAAAACCCAAATCCGTCACTAACATTTTAAACCAACCCATGTTCCA 180
DB 131 CAGTGGATGTCAAAACCCAAATCCGTCACTAACATTTTAAACCAACCCATGTTCCCT 190
QY 181 AAATTGGGGGTGATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCTTCACCGGTGT 240
DB 191 AAATTAGGGGTAAATGCTTGTGGGTGGGGTGGAAACAAAGGCTTACCTTCACCGGTGT 250
QY 241 GTTATTGTCTAAGAGAGACATTTTCATGGGCTACAAAGAGCAAGATTCAACAGCCAT 300
DB 251 GTTATTGTCTAAGAGAGGCACTTTTCATGGGCTACAAAGAGCAAGATTCAACAGCCAT 310
QY 301 TACCTTGGCTCCCTCAACCCAGCCTCAAGTATTCAGATTGATCTTCCAGGAGAGAA 360
DB 311 TACCTTGGCTCCCTCAACCCAGCCTCAAGTATTCAGATTGATCTTCCAGGAGAGAA 370
QY 361 ATCTATGCCCCCTTCAAGAGTCTGCTTCCAATGGTTAACTTCAGAGCACTTGTGTTGGG 420
DB 371 ATCTATGCCCCCTTCAAGAGCCTGCTTCCAATGGTTAACTTCAGAGCACTTGTGTTGGG 430
QY 421 GGATGGGATATGAGCAATGAACTGGCTGATGCCATGGCCAGGCAAAAGGTGTTTAC 480
DB 431 GGATGGGATATGAGCAATGAACTGGCTGATGCCATGGCCAGGCAAAAGGTGTTTAC 490
QY 481 ATCGATTTTGCAGAGCAGTTGAGGCTTTAATGGAATTCATGATGCTTCCCGGAATC 540
DB 491 ATCGATTTTGCAGAGCAGTTGAGGCTTTAATGGAATTCATGATGCTTCCCGGAATC 550
QY 541 TACGACCCCGGATTTTCTTGTCTGCCAACAGAGAGCGTGGCCAAACAGTGAATTAAGGCG 600
DB 551 TATGACCCCGGATTTTCTTGTCTGCCAACAGAGAGCGTGGCCAAACAGTGAATTAAGGCG 610
QY 601 ACAAAGCAAGAGCAAGTTCAGCAAAATCATCAAGAGCATCAAGCGCTTTAAGGAAGCCACC 660
DB 611 ACAAAGCAAGAGCAAGTTCAGCAAAATCATCAAGAGCATCAAGCGCTTTAAGGAAGCCACC 670
QY 661 AAAGTGACAGAGGTGTTGCTCTGTGACTGCCAACAGAGAGGTATAGCAATTTGGTT 720
DB 671 AAAGTGACAGAGGTGTTGCTCTGTGACTGCCAACAGAGAGGTATAGCAATTTGGTT 730
QY 721 GTAGGCTTAAATGACCACTGAGAAATCTTGGCTGCTGTGAGACAGAAATGAGGTGAG 780
DB 731 GTAGGCTTAAATGACCACTGAGAAATCTTGGCTGCTGTGAGACAGAAATGAGGTGAG 790
QY 781 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGTATGGAATAATGTTCTTTTCAATTAAT 840
DB 791 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGTATGGAATAATGTTCTTTTCAATTAAT 850
QY 841 GGAAGCCCTCAGAAACCTTTTGTACCAAGGCGTGAATGATCTTGCCATTCGCGAGAACT 900
DB 851 GGAAGCCCTCAGAAACCTTTTGTACCAAGGCGTGAATGATCTTGCCATTCGCGAGAACT 910

QY 901 TTGATTTGGAGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGGTTGAT 960
DB 911 TTGATTTGGAGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGGTTGAT 970
QY 961 TTTCTTGTGGGGGCTGTGATCAAGCCAACTCTATAGTTAGTTTCAACCACTCTGGGAAC 1020
DB 971 TTTCTTGTGGGGGCTGTGATCAAGCCAACTCTATAGTTAGTTTCAACCACTCTGGGAAC 1030
QY 1021 AATGATGATGATCTCTCGGCTTCACAAACCTTCGCTCCAGAAATCTCCAGAGC 1080
DB 1031 AATGATGATGATCTCTCGGCTTCACAAACCTTCGCTCCAGAAATCTCCAGAGC 1090
QY 1081 AAGCTTTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1091 AAGCTTTGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1150
QY 1141 GACCATGTTGTTGTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1151 GACCATGTTGTTGTTATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1210
QY 1201 GAGTACACTTCAAGATATTTCAAGGCTGGAAGAAACACCATTTTTCACAAACATGAT 1260
DB 1211 GAGTACACTTCAAGATATTTCAAGGCTGGAAGAAACACCATTTTTCACAAACATGAT 1270
QY 1261 GAGATTCCTCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1271 GAGATTCCTCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1330
QY 1321 ACTAGAAATCAAGTTTAAAGCTGAAATGAGGAAATTTCACTCATTCACCCAGTTGCT 1380
DB 1331 ACTAGAAATCAAGTTTAAAGCTGAAATGAGGAAATTTCACTCATTCACCCAGTTGCT 1390
QY 1381 ACCATTTCCAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACCAAGTGTGAT 1440
DB 1391 ACCATTTCCAGCTATCTGACCAAGGCTCTCTGTTTCCACCGGGTACCAAGTGTGAT 1450
QY 1441 GCATTGTCAAGAGCAGCTGCAATGCTGGAATAATAGAGGCTTGTGTTGATTTGGCC 1500
DB 1451 GCATTGTCAAGAGCAGCTGCAATGCTGGAATAATAGAGGCTTGTGTTGATTTGGCC 1510
QY 1501 CCAGAAATTAACATGATTTCTGAGTACAAAGTGA 1533
DB 1511 CCAGAAATTAACATGATTTCTGAGTACAAAGTGA 1543

RESULT 2
BD075266 1782 bp DNA linear PAT 27-AUG-2002
LOCUS BD075266
DEFINITION Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid.
ACCESSION BD075266
VERSION BD075266.1 GI:22620869
KEYWORDS JP 2001519665-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1782)
AUTHORS Hitz,W.D. and Sebastian,S.A.
TITLES Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid
JOURNAL Patent: JP 2001519665-A 1 23-OCT-2001;
EI DU PONT DE NEMOURS AND CO
COMMENT OS Soybean line LR13
PN JP 2001519665-A/1
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN
PC CI2NI5/52, CI2NI5/82, CI2NI5/11, CI2NI5/90, A01H5/00 CC
Strandness: Double;
CC Topology: linear;
CC Soybean plant producing seeds with reduced

Levels of raffinose
CC and phytic acid Location/Qualifiers
FH Key 54.1586.
FT CDS Location/Qualifiers
1.1782
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 95.5%; Score 1464.2; DB 6; Length 1782;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

1 ATGTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAAGTACCCGAGCTGAGATT 60
54 ATGTCATGAGAAATTTTAAAGTAGAGTCTTAATGTGAAGTACCCGAGCTGAGATT 113
61 CAGTCCGTGTACACACGAAACACCCGAACTTTTCAAGAAACAGAAATGACCTTAT 120
114 CAGTCCGTGTACACACGAAACACCCGAACTTTTCAAGAAACAGAAATGACCTTAT 173
121 CAGTGGATGTGCAAAACCAATCCGTCACTACCAATTTTAAACCAACCAATGTTTCA 180
174 CAGTGGATGTGCAAAACCAATCCGTCACTACCAATTTTAAACCAACCAATGTTTCT 233
181 AAATTTGGGGGTATGCTTGGGGTGGGGTGAACCAACGCTCTACCCCTCAACCGGTGT 240
224 AAATTTGGGGGTATGCTTGGGGTGGGGTGAACCAACGCTCTACCCCTCAACCGGTGT 293
241 GTTATTGTCTAACAGAGACATTTTATGTGGCTTCAAGAGACAGATTCAACAGCAAT 300
294 GTTATTGTCTAACAGAGACATTTTATGTGGCTTCAAGAGACAGATTCAACAGCAAT 353
301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTTCCAGTTGAGTTCCTTCCAGGAGAGAA 360
354 TACTTTGGCTCCCTCAACCAAGCTCAGCTATTTCCAGTTGAGTTCCTTCCAGGAGAGAA 413
361 ATCTATGCCCATTTAGAGAGTGTGCTTCCATGTGTTTATCTGACGACATTTGTTGGG 420
414 ATCTATGCCCATTTAGAGAGTGTGCTTCCATGTGTTTATCTGACGACATTTGTTGGG 473
421 GGATGGGATATGACAACTGAACCTGCTGATGCCATGCGCAAGGCAAAAGTGTTCAC 480
474 GGATGGGATATGACAACTGAACCTGCTGATGCCATGCGCAAGGCAAAAGTGTTCAC 533
481 ATCGATTTGACAGACAGTTGAGGCTTACATGGAATCCATGCTCCCGGAAATC 540
534 ATCGATTTGACAGACAGTTGAGGCTTACATGGAATCCATGCTCCCGGAAATC 593
541 TAGACCCCGGATTTTATGCTGCTCCAAACCAAGAGAGCGTCCAAACAGTGAATTAAGGC 600
594 TATGACCCCGGATTTTATGCTGCTCCAAACCAAGAGAGCGTCCAAACAGTGAATTAAGGC 653
601 ACAAAGCAAGAGCAAGTTGACGAAATCATCAAAAGCATCAAGCGCTTTAAGAACCCACC 660
654 ACAAAGCAAGAGCAAGTTGACGAAATCATCAAAAGCATCAAGCGCTTTAAGAACCCACC 713
661 AAAGTGACCAAGGTTGCTGCTGAGACTGCCAACAGAGAGTTATGAAATTTTGGTT 720
714 AAAGTGACCAAGGTTGCTGCTGAGACTGCCAACAGAGAGTTATGAAATTTTGGTT 773
721 GTAGGCTTTAATGACACATGAGAAATCTTTGGCTGTGTGAGACAGAAATGAGGCTGAG 780
774 GTAGGCTTTAATGACACATGAGAAATCTTTGGCTGTGTGAGACAGAAATGAGGCTGAG 833
781 ATTTTCTCTTCACTTTGATGCAATGCTGTGTGATGAAATGTTCTTTCAATTAAT 840
834 ATTTTCTCTTCACTTTGATGCAATGCTGTGTGATGAAATGTTCTTTCAATTAAT 893
841 GGAAGCCCTCAGAAACATTTTGTACAGGGCTGATTTGATCTTGCCATCGCGAGAAACAT 900

|||||
894 GGAAGCCCTCAGAAACATTTTGTACAGGGCTGATTTGATCTTGCCATCCCGAGAAACAT 953
901 TTGATTTGTGAGAGATGACTTTCAAGAGTGTGACAGCCAAATGAATCTGTGTTGTTGAT 960
954 TTGATTTGTGAGAGATGACTTTCAAGAGTGTGACAGCCAAATGAATCTGTGTTGTTGAT 1013
961 TTTTCTTTGGGGGCTGTGATCAAGCCAACTTATATGTTAGTTTCAACCAATCTGGGAAAC 1020
1014 TTCTTTGGGGGCTGTGATCAAGCCAACTTATATGTTAGTTTCAACCAATCTGGGAAAC 1073
1021 AATGATGTGATGATCTCTCGGCTCCAAACCTTCCGCTCCAAAGAAATCTCAAGAGC 1080
1074 AATGATGTGATGATCTTTCGCTCCAAACCTTTCGTTCCAAAGAAATCTCAAGAGC 1133
1081 AACGTTGTTGAGATATGTTTCAAGAGATGCAATGCAATGCTTATGAGGCTGTGAAATCCA 1140
1134 AACGTTGTTGAGATATGTTTCAAGAGATGCAATGCAATGCTTATGAGGCTGTGAAATCCA 1193
1141 GACCATGTTGTTGTTATTTAATGATGCTTACGTTAGGGGATAGCAAGAGACCATGAT 1200
1194 GACCATGTTGTTGTTATTTAATGATGCTTACGTTAGGGGATAGCAAGAGACCATGAT 1253
1201 GAGTACACTTCAAGATATTTCAAGAGATGCAATGCAATGCTTATGAGGCTGTGAAATCCA 1260
1254 GAGTACACTTCAAGATATTTCAAGAGATGCAATGCAATGCTTATGAGGCTGTGAAATCCA 1313
1261 GAGATTTCCCTTTTAAAGCTGCTCTTATTTATCTTGAATGTTGCTTCTTGTGAGCTGAC 1320
1314 GAGATTTCCCTTTTAAAGCTGCTCTTATTTATCTTGAATGTTGCTTCTTGTGAGCTGAC 1373
1321 ACTAAGATTCAGTTTAAAGCTGCTCTTATTTATCTTGAATGTTGCTTCTTGTGAGCTGAC 1380
1374 ACTAAGATTCAGTTTAAAGCTGCTCTTATTTATCTTGAATGTTGCTTCTTGTGAGCTGAC 1433
1381 ACCATTTCCAGCTATCTGACCAAGGCTCCTGTTTCAACCGGGTACCAAGTGTGAT 1440
1434 ACCATTTCCAGCTATCTGACCAAGGCTCCTGTTTCAACCGGGTACCAAGTGTGAT 1493
1441 GCATTTGCAAAAGCAGCTGCTGATGCTGTAAGAAACATATATGAGGCTTGTGTTGATGGCC 1500
1494 GCATTTGCAAAAGCAGCTGCTGATGCTGTAAGAAACATATATGAGGCTTGTGTTGATGGCC 1553
1501 CCAAGAAATACATGATTTCTGAGTTCAAGTGA 1533
1554 CCAAGAAATACATGATTTCTGAGTTCAAGTGA 1586

DB
RESULT 3
BD075269 1533 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION
Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid.
ACCSSION
BD075269
VERSION
BD075269.1 GI:22620872
KEYWORDS
JP 2001519665-A/4.
SOURCE
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 1533)
Hitz,W.D. and Sebastian,S.A.
Soybean plant producing seeds with reduced levels of raffinose
saccharides and phytic acid
Patent: JP 2001519665-A 4 23-OCT-2001;
EI DU PONT DE NEMOURS AND CO
OS Soybean line UR33
PN JP 2001519665-A/4
PD 23-OCT-2001
PF 07-APR-1998 JP 1998543012
PR 08-APR-1997 US 08/835751
PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN
PC C12N15/52,C12N15/82,C12N15/11,C12N9/90,A01H5/00 CC
Strandedness: Double;

CC Topology: Linear;
 CC Soybean plant producing seeds with reduced
 levels of raffinose
 CC and phytic acid saccharides
 FH Key Location/Qualifiers
 FT CDS 1..1533.

FEATURES
 source 1..1533
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN

Query Match 95.4%; Score 1462.6; DB 6; Length 1533;
 Best Local Similarity 97.1%; Pred. No. 0;
 Matches 1489; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```

QY 1 ATGTTTCATCGAGAAATTTTAAAGTAGAGATCCCTAAATGTGAATAGACCGAGACTGAGATT 60
DB 1 ATGTTTCATCGAGAAATTTTAAAGTAGAGATCCCTAAATGTGAATAGACCGAGACTGAGATT 60
QY 61 CAGTCCTGTGTACACTAGAGAAACACCGAATTTTTCAGAGAAACAGGAATGGCACTAT 120
DB 61 CAGTCCTGTGTACACTAGAGAAACACCGAATTTTTCAGAGAAACAGGAATGGCACTAT 120
QY 121 CAGTGGATTTGTCAAAACCCAAATCCGTCAACTCAATTTAAACCAACCCATGTTCCA 180
DB 121 CAGTGGATTTGTCAAAACCCAAATCCGTCAACTCAATTTAAACCAACCTCATGTTTCT 180
QY 181 AAATTTGGGGGTGATCTGTGGGTTGGGGTGGAAACACGGCTCTACCCCTCAACGGTGGT 240
DB 181 AAATTTGGGGGTGATCTGTGGGTTGGGGTGGAAACACGGCTCTACCCCTCAACGGTGGT 240
QY 241 GTTATTTGCTTAAACAGAGAGACATTTTCAATGGGCTACAAAGAGACAAATTTCAACAGCCAT 300
DB 241 GTTATTTGCTTAAACAGAGAGACATTTTCAATGGGCTACAAAGAGACAAATTTCAACAGCCAT 300
QY 301 TACTTTGGGCTCCCTACCCCAACCTCAGCTATTTCAGATTGATTCCTTCCAGGAGAGAGAA 360
DB 301 TACTTTGGGCTCCCTACCCCAACCTCAGCTATTTCAGATTGATTCCTTCCAGGAGAGAGAA 360
QY 361 ATCTATGCCCATTCAGAGAGTCTGCTTCCAAATGTTAATCTCGACGACATTTGTTTGGG 420
DB 361 ATCTATGCCCATTCAGAGAGTCTGCTTCCAAATGTTAATCTCGACGACATTTGTTTGGG 420
QY 421 GGATGGGATATACGACAACTGAACCTGGCTGATGCCATGGCCCAAGGCAAAAGTGTGGAC 480
DB 421 GGATGGGATATACGACAACTGAACCTGGCTGATGCCATGGCCCAAGGCAAAAGTGTGGAC 480
QY 481 ATCGATTTGCAGAGAGAGTGAAGGCTTACATGGAATCGATGGTTCACCTCCCGGAATC 540
DB 481 ATCGATTTGCAGAGAGAGTGAAGGCTTACATGGAATCGATGGTTCACCTCCCGGAATC 540
QY 541 TACGACCCCGGATTTTCAATGCTGCCAACCAAGAGAGCGCTGCCAACCAACGTGATTAGGGC 600
DB 541 TACGACCCCGGATTTTCAATGCTGCCAACCAAGAGAGCGCTGCCAACCAACGTGATTAGGGC 600
QY 541 TATGACCCGGATTTTCAATGCTGCCAACCAAGAGAGCGCTGCCAACCAACGTGATTAGGGC 600
DB 541 TATGACCCGGATTTTCAATGCTGCCAACCAAGAGAGCGCTGCCAACCAACGTGATTAGGGC 600
QY 601 ACAAAGCAAGACCAAGTTCAGCAATCATCAAAAGACATCAAGCGCTTTAAGAGACACCC 660
DB 601 ACAAAGCAAGACCAAGTTCAGCAATCATCAAAAGACATCAAGCGCTTTAAGAGACACCC 660
QY 661 AAAGTGAACAAGGTGTGTCTGCTGAGCTGCCAACAGAGAGGTATGCAATTTGGTT 720
DB 661 AAAGTGAACAAGGTGTGTCTGCTGAGCTGCCAACAGAGAGGTATGCAATTTGGTT 720
QY 721 TAGAGCTTAAATGACACATGAGATCTTTGGCTGTGTGAGACAGAAATGAGGCTGAG 780
DB 721 TAGAGCTTAAATGACACATGAGATCTTTGGCTGTGTGAGACAGAAATGAGGCTGAG 780
QY 781 ATTTCTCTTCCACCTTGTATGCTGCTGTGTGATGGAATAATTTCTTTCAATTAAT 840
DB 781 ATTTCTCTTCCACCTTGTATGCTGCTGTGTGATGGAATAATTTCTTTCAATTAAT 840

```

```

QY 841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATGTAATCTTGGCATCGGAGAAACT 900
DB 841 GGAAGCCCTCAGAACACTTTTGTACCAAGGCTGATGTAATCTTGGCATCGGAGAAACT 900
QY 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAATATCTGTGTGGTAT 960
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGCAAAATGAATATCTGTGTGGTAT 960
QY 961 TTTCTGTGGGGGCTGGTATCAGCAACATCTATAGTATGTTACCAACCATCTGGGAAAC 1020
DB 961 TTTCTGTGGGGGCTGGTATCAGCAACATCTATAGTATGTTACCAACCATCTGGGAAAC 1020
QY 1021 AATGATGATGAATCTCTCGGCTTCCAAACCTTCGCTCCAAAGAAATCTCCAAGAC 1080
DB 1021 AATGATGATGAATCTCTCGGCTTCCAAACCTTCGCTCCAAAGAAATCTCCAAGAC 1080
QY 1081 AACGTTTGTGACATATGCTCAACGACATGCCATCTCTATGAGCCCTGTGAAACATCCC 1140
DB 1081 AACGTTTGTGATGATATGCTCAACGACATGCCATCTCTATGAGCCCTGTGAAACATCCA 1140
QY 1141 GACCATGTTGTTGTTATTAATGATGTCCTTACGTAAGGAGATGACAGAGACCATGAT 1200
DB 1141 GACCATGTTGTTGTTATTAATGATGTCCTTACGTAAGGAGATGACAGAGACCATGAT 1200
QY 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCATTTGTTTGCACAAACATGT 1260
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAACACCATTTGTTTGCACAAACATGT 1260
QY 1261 GAGCATTCCTTTTATAGCTGCTCTATTAATCTTGAACCTTGCTCTTCTTGTGAGCTGAG 1320
DB 1261 GAGCATTCCTTTTATAGCTGCTCTATTAATCTTGAACCTTGCTCTTCTTGTGAGCTGAG 1320
QY 1321 ACTAGATTCAGTTTAAAGCTGAAATAAGAGGAAATTCACATCTCCACCCAGTTGCT 1380
DB 1321 ACTAGATTCAGTTTAAAGCTGAAATAAGAGGAAATTCACATCTCCACCCAGTTGCT 1380
QY 1381 ACCATTCCTCAGCTATCTGACCAAGGCTCTCTGTTCCACCGGCTACCAAGTGTGAT 1440
DB 1381 ACCATTCCTCAGCTATCTGACCAAGGCTCTCTGTTCCACCGGCTACCAAGTGTGAT 1440
QY 1441 GCATTGTCAAAAGCAGCTGCAATGCTGAAAAACATTAAGAGGCTGTGTGATTGACC 1500
DB 1441 GCATTGTCAAAAGCAGCTGCAATGCTGAAAAACATTAAGAGGCTGTGTGATTGACC 1500
QY 1501 CCAGAAATTAACATGATCTCGAGTACAAAGTGA 1533
DB 1501 CCAGAAATTAACATGATCTCGAGTACAAAGTGA 1533

```

RESULT 4
 AF293970 1791 bp mRNA linear PLN 03-MAY-2001
 LOCUS Glycine max myo-inositol-3-phosphate synthase (MIRS) mRNA, complete
 DEFINITION cds.
 ACCESSION AF293970
 VERSION AF293970.1 GI:13936690
 KEYWORDS
 SOURCE Glycine max (soybean)
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 REFERENCE 1 (bases 1 to 1791)
 AUTHORS Hegeman, C.E., Good, L.L. and Grabau, E.A.
 TITLE Expression of D-myo-inositol-3-phosphate synthase in soybean.
 JOURNAL Implications for phytic acid biosynthesis
 Plant Physiol. 125 (4), 1941-1948 (2001)
 MEDLINE 21196082
 PUBMED 1129373
 REFERENCE 2 (bases 1 to 1791)
 AUTHORS Hegeman, C.E., Good, L.L. and Grabau, E.A.

TITLE Direct Submission
JOURNAL Submitted (07-AUG-2000) Plant Pathology, Physiology, and Weed
Science, Virginia Tech, Fralin Biotechnology Center, Blacksburg, VA
24060, USA

FEATURES
source Location/Qualifiers
1..1791
/organism="Glycine max"
/mol_type="mRNA"
/culivar="Williams 82"
/db_xref="taxon:3847"
/dev_stage="Immature cotyledons"
1..1791
/gene="MIPS"
62..1594
/gene="MIPS"
/note="Isomerase"
/codon_start=1
/product="myo-inositol-3-phosphate synthase"
/protein_id="AAK49896.1"
/db_xref="GI:13936691"

gene
CDS

ORIGIN
Query Match 95.0%; Score 1456.2; DB 8; Length 1791;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1485; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
1 ATGTTTCATGAGAAATTTTAAAGTAGAGAGTCTTAATGTAGATACACGAGACTGAGATT 60
62 ATGTTTCATGAGAAATTTTAAAGTAGAGTCTTAATGTAGATACACGAGACTGAGATT 121
61 CAGTCGGTGTACAACTACGAAACCAACGAACTTGTTCACGAGAACGAGATGGCACTAT 120
122 CAGTCGGTGTACAACTACGAAACCAACGAACTTGTTCACGAGAACGAGATGGCACTAT 181
121 CAGTCGATTTGTCAAAACCAATCTGCTCAACTACCAATTTTAAACCAACCACTGTTCCA 180
182 CAGTCGATTTGTCAAAACCAATCTGCTCAACTACCAATTTTAAACCAACCACTGTTCCA 241
181 AAATTGGGGGTGATGTTGGGTTGGGGTGGAAACAAGGCTTACCGTCAACCGGTGGT 240
242 AAATTGGGGGTGATGTTGGGTTGGGGTGGAAACAAGGCTTACCGTCAACCGGTGGT 301
241 GTTATTGTCTAACAGAGAGACATTTTCATGGGCTAACAAAGACAGATTTCAACAAGCAAT 300
302 GTTATTGTCTAACAGAGAGGCACTTTTCATGGGCTAACAAAGACAGATTTCAACAAGCAAT 361
301 TACTTTGGCTCCCTCAACCAAGCTCTCAAGTATTCAGATTGATCTTCCAGGGAAGAA 360
362 TACTTTGGCTCCCTCAACCAAGCTCTCAAGTATTCAGATTGATCTTCCAGGGAAGAA 421
361 ATCTATGCCCATTTAAGAGTCTGCTTCAATGGTTAATCTTGAAGACATTTGTGTTGGG 420
422 ATCTATGCCCATTTAAGAGCTGCTTCAATGGTTAATCTTGAAGACATTTGTGTTGGG 481
421 GGATGGGATATCAGCAACATGAACCTGGCTGATGTCATGGCAGGCAAAAGGTGTTGAC 480
482 GGATGGGATATCAGCAACATGAACCTGGCTGATGTCATGGCAGGCAAAAGGTGTTGAC 541
481 ATCGATTTCAGAGACAGATTGAGGCTTATCATGAAATCATGATCTCCCGGAATC 540
542 ATCGATTTCAGAGACAGATTGAGGCTTATCATGAAATCATGATCTCCCGGAATC 601
541 TAGCAACCGGATTTATGCTGCTCAACCAAGAGGCGTGCACAACATGATTAAAGGC 600

|||||
602 TATGACCCGGATTTCACTTCTGCTCCAAAGAGAGCGTCCAAACACTCATCAAGGCC 661
601 ACAAAGCAAGACCAAGTTTCAAGAAATCATCAAGACATCAAGCGTTTAAAGAGCCACC 660
662 ACAAAGCAAGACCAAGTTTCAAGAAATCATCAAGACATCAAGCGTTTAAAGAGCCACC 721
661 AAGTGGACAAAGTGTGTTCTCTGTGACTGGCAACACAGAGAGATATGCAATTTGGTT 720
722 AAGTGGACAAAGTGTGTTCTCTGTGACTGGCAACACAGAGAGATATGCAATTTGGTT 781
721 GTAGGCTTAAATGACACCATGAGAAATCTTGGCTGCTGTGGAAGAAATGAGGCTGAG 780
782 GTGGGCTTAAATGACACCATGAGAAATCTTGGCTGCTGTGGAAGAAATGAGGCTGAG 841
842 ATTTCTCTTCCACCTTGTATGCCATTTGCTGTGTAAGAAATGTTCTTTCAATTAAT 840
841 GGAGCCCTCAGAACACTTTTGTACAGGGCTGATTTGATCTTGGCATGGCAGAACACT 900
902 GGAGCCCTCAGAACACTTTTGTACAGGGCTGATTTGATCTTGGCATGGCAGAACACT 961
901 TTGATTGGTGGAGATGACTTCAAGAGTGGTCAAGCAAAATGAAATCTGTGTTGGTTAT 960
962 TTGATTGGTGGAGATGATTTCAAGAGTGGTCAAGCAAAATGAAATCTGTGTTGGTTAT 1021
961 TTTCTTGGGGGGCTGTATCAAGCAACATCATATGTTAGTTACCAACCATCTGGGAAC 1020
1022 TTTCTTGGGGGGCTGTATCAAGCAACATCATATGTTAGTTACCAACCATCTGGGAAC 1081
1021 AATGATGATATGATCTCTGGCTTCAACAACTCCGCTCAAGAAATCTTCAAGAGC 1080
1082 AATGATGATATGATCTTGGCTTCAACAACTCCGCTTCAAGAAATCTTCAAGAGC 1141
1081 AAGCTTGTGAGATGATGATCAAGCAATGCAATGCTTATGAGCGTGGTGAACATCCC 1140
1142 AAGCTTGTGAGATGATGATCAAGCAATGCAATGCTTATGAGCGTGGTGAACATCCA 1201
1141 GACCATGTTGTTTATTAAGTATGCTTACGTAAGGGATAGCAAGAGCCATGAT 1200
1202 GACCATGTTGTTTATTAAGTATGCTTACGTAAGGGATAGCAAGAGCCATGAT 1261
1201 GAGTACACTTCAAGATATTTCAAGGTTGGAAGAACACCATTTGTTTGAACAACATGT 1260
1262 GAGTACACTTCAAGATATTTCAAGGTTGGAAGAACACCATTTGTTTGAACAACATGC 1321
1261 GAGATATCCCTTTTGAAGTCTCTTATTAATCTTGAAGTCTGCTTCTTGGCTGAGCGT 1320
1322 GAGATATCCCTTTTGAAGTCTCTTATTAATCTTGAAGTCTGCTTCTTGGCTGAGCGT 1381
1321 ACTAGATATCCAGTTTAAAGCTGGAAGTGAAGGAAATTCATCTCAATCCACAGTTGCT 1380
1382 ACTAGATATCCAGTTTAAAGCTGGAAGTGAAGGAAATTCATCTCAATCCACAGTTGCT 1441
1381 ACCATTCTCAGCTATCTCAACCAAGGCTCTCTGTTTCAACCGGATCAACAGTGTGAT 1440
1442 ACCATTCTCAGCTATCTCAACCAAGGCTCTCTGTTTCAACCGGATCAACAGTGTGAT 1501
1441 GCATTGTCAAGACACCGTGAATGCTGGAAGAAATTAAGAGGCTGTGTTGATTTGGCC 1500
1502 GCATTGTCAAGACACCGTGAATGCTGGAAGAAATTAAGAGGCTGTGTTGATTTGGCC 1561
1501 CCAGAGAAATTAATGATTTCTGAGATCAAGTGA 1533
1562 CCAGAGAAATTAATGATTTCTGAGATCAAGTGA 1594

RESULT 5
AB032073 1931 bp mRNA linear PLN 08-SEP-1999
LOCUS Nicotiana glauca Nginp51 mRNA for myo-inositol-1-phosphate
DEFINITION synthase, complete cds.

ACCESSION AB032073
VERSION AB032073.1 GI:5834499
KEYWORDS myo-inositol-1-phosphate synthase.
SOURCE Nicotiana paniculata
ORGANISM Nicotiana paniculata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 (bases 1 to 1931)
AUTHORS Hashimoto, A., Yamada, S. and Komori, T.
TITLE Myo-inositol-1-phosphate synthase
JOURNAL Published Only in Database (1999)
AUTHORS 2 (bases 1 to 1931)
TITLE Hashimoto, A., Yamada, S. and Komori, T.
JOURNAL Direct Submission
Submitted (01-SEP-1999) Akiko Hashimoto, Japan Tobacco Inc., Plant
Breeding and Genetics Research Laboratory, 700 Higashibara, Iwata,
Toyoda-cho, Shizuoka 438-0802, Japan
(E-mail: Akiko.Hashimoto@pbgrl.jti.co.jp, Tel: +81-538-32-7116,
Fax: +81-538-33-6046)
FEATURES
source 1. 1931
Location/Qualifiers
/organism="Nicotiana paniculata"
/mol_type="mRNA"
/db_xref="taxon:62141"
1. 1931
/gene="NPINP1"
92. 1624
/gene="NPINP1"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="BAA84084.1"
/db_xref="GI:5834500"
/translation="MEIENFKYSEPNVKTESEIHSYDYOTTELVHDEKNGTYOMTV
KPKTYKBEKDVHVKLGVMLVGMVGNVGLTGVLANREGISLMAKDKYQKANYE
GSLTASTIRVSENGEBIYAPFKSLIPRVNDDVYFGMDLSDMLADMAKRAYPD
IDLQKLRYESWVPLPGITDPDPIAAGQSRANNVKGTKEDIDQIKDIREFKE
KNKVDVNLWANTERISNVVGLNDTMEMLFASVDREAEISSTLYALACILEEN
PFINSPQTFVPLGIDLAIKKNTLIGDDPFKQTKMSVLDPLVAGIKPISIV
YVHLNNGNMLSAPOTPRSKEISKSNVDDVSSNALIYEGSEHDPHVYKVPYV
GDSKRAMYEYSEIFMGKNTLVHNTGDSILAAPIILDVLAELSRITOLKEGE
GKFSPHPVATILSTIKAPLVPGTPPVNALSKORAMLENTLRACVGLAPENNMLE
YK"

ORIGIN
Query Match 71.6%; Score 1097.8; DB 8; Length 1931;
Best Local Similarity 82.3%; Pred. No. 8.3e-289;
Matches 1261; Conservative 0; Mismatches 272; Indels 0; Gaps 0;
1 ATGTTTCATCGAGAAATTTTAAAGTAGAGAGTCTTAATGTGAAGTACACCGAGACTGAGATT 60
2 AATGTTTATGAGAAATTTTAAAGTTGAGAGCCCAAGTTAAGTACCGAAGATGMAATT 151
61 CAGTCCGTTTACAACTACGAAACCACCGAATTGTTTACGAGAAACGAAATGCGACTTAT 120
152 CACTCTGTCTATGATTATCAAACTGAGTTAGTTTCATGATGAGAAAATGAGCAATAT 211
121 CAGTGAATTTGCAAAACCAATCCGTCATCACTAATTTAAACCAACACCAATGTTTCA 180
212 CAATGACCGGTCAAGCTTAAGACTGTCAATATGATGATTAAGACTATGTTCAATGTTCCC 271
181 AAATGGGGGGTATGCTTGTGGGTGGGGTGGAAACAAGGCTTACCCCTCAACCGGTGCT 240
272 AATTTAGGGGTTATGCTTGTGGATGGGGTGGAAAACAATGTTCAACCTTTGACCGGTGCT 331
241 GTTATTGTTCAACAGAGACATTTTATGAGGCTTCAAAAGACAAAGTTCAAACAGCAAT 300
332 GTTATTGTTCAACAGAGAAATTTTATGAGGCTTCAAAAGATTAAGTGGCAACAGCAAT 391
301 TACTTTGGCTCCCTCAACCAAGCTCAGGCTATTGAGTTGAGTCTTCCAGGAGAGAGAA 360
392 TACTTTGGCTCTCTTAAGCTTCAAGGCTTCACTATTGAGTTGGGCTTTTCAATGAGAGAG 451

361 ATCTATGCCCATTCAGAGAGTCTGTTCCATGTTAATCTGTACAGCATTTGTTGGG 420
452 ATCTATGCTCCATTTTAAAGCCTCTTCCATGTGTCATTCAGATGAGTGTGTTGGA 511
421 GGATGGGATATCAGCAATGAAACCTGGCTGATGTCATGAGGCGACAGGCAAGTGTGTTGAC 480
512 GGATGGGATATCAGCAATGAAATTTTACAGATGCCATGCGCAGGGCTTAAGTATTTGAT 571
481 ATCGATTTGCAAGACAGTGGAGCTTACATGAAATCATGTTTCCACTCCCGGAATC 540
572 ATGATCTACAAAAGCAATGTAGAGGCTTACATGGAATCTATGTTCCACTCCCTGATTC 631
541 TAGACCCGGATTTTATGCTGCTCCCAACCAAGAGAGCTGCCACACAGTATTAAGGC 600
632 TAGACCCGTGATTTTATGCTGCTTAAACCAAGGTCACCTGGCAACAGTATTAAGGA 691
601 ACAAGCAAGAGCAAGTTCAGCAATCATCAAGACATCAAGGCGTTTAAAGAACCCACC 660
692 ACCAAGAAAGAACCAATGATCAATCATTAAGATATTAGGAGTTTAAAGAAAAGAAC 751
661 AAAGTGACAAAGTGTGTTGCTGCTGAGCTGCCAACAAGAGGTATAGCAATTTGGTT 720
752 AAAGTGACAAAGTGTGTTGTTGAGCTGCTAACACTGAAGATACAGTATGTTGGTT 811
721 GTAGGCTTTATGACACATGAGACATCTTTGGCTGCTGTGAGACAGAAATGAGGCTGAG 780
812 GTTGACCTTAAATGACACTATGAGAAAACCTTTTGGCTGTGAGACAGAAATGAACTGAA 871
781 ATTTCTCTTCCACCTGTGATGCCATGCTGCTGTGATGAGAAAATGTTCTCTTCTTAAT 840
872 ATATCTCTTCCACTTTGATGCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTTATTAAT 931
841 GGAAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTGGCATGCGAGAGAACT 900
932 GGAAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTGGCATGCGAGAGAACT 991
901 TTGATTTGAGATGATGATCTTCAAGAGTGTGAGACCAAAATGAATCTGTGTTGGTAT 960
992 TTGATTTGAGATGATGATCTTCAAGAGTGTGAGACCAAAATGAATCTGTGTTGGTAT 1051
961 TTTCTTGGGGGGCTGATATCAAGCAACATCTATGATGATGATGATGATGATGATGATGAT 1020
1052 TTCTTGTGAGAGCTGTGATTAAGCAACATCAATGATGATGATGATGATGATGATGATGAT 1111
1021 AATGATGATGATGATCTGCTGCTTCCAAACCTTCCGCTCCAAAGAAATCTTCAAGAGC 1080
1112 AATGATGATGATGATCTGCTGCTTCCAAACCTTCCGCTCCAAAGAAATCTTCAAGAGC 1171
1081 AAGTTGTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1172 AATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
1141 GACGATGTTGTTGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1232 GACGATGTTGTTGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1291
1201 GAGTACCTTCAAGATATTTCAAGGTTGAGAAAGAACACATTTGTTTCAACAACATGTT 1260
1292 GAGTACATCTGAGATTTTCAAGGTTGAGAAAGAACACATTTGTTTCAACAACATGTT 1351
1261 GAGATTCCTTTTATGCTGCTCTTATTTATCTTGAATCTTGTCTTGTCTGAGTGAAGC 1320
1352 GAGATTTCACTTTTATGCTGCTCTTATTTATGATGATGATGATGATGATGATGATGAT 1411
1321 ACTAAGATCAGTTTAAAGCTGAGAAATGAGGAGAAATTCATCTATTCACACCACTGCT 1380
1412 ACCCGATTCACCTCAAGCTGAGAAAGGAGGAGGTTAGTTTCACTCTTTCACACCCGTTGCT 1471
1381 ACCATTTCTAGCTATCTGACCAAGGCTCTCTGTTTCAACCGGATGACACCACTGTTGAT 1440
1472 ACTATCTCAGTACCTTACCAAGGCTCTCTGTTTCAACCGGATGACACCACTGTTGAT 1531
1441 GCATTGTCAAGACGCTGCAATGCTGAGAAAACATTAAGAGGCTTGTGTTGATTTGGCC 1500

Db 1532 GCACTCTCAAGAGAGGGGCAATGCTTGAGAACATTTGAGGGCTTGCTGAGACTTGCA 1591
Qy 1501 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1533
Db 1592 CCAGAGAACACATGATTTCTGGAATACAAATGA 1624

RESULT 6
E27176
LOCUS E27176 1950 bp DNA linear PAT 18-JUN-2001
DEFINITION Novel INPS gene derived from nicotiana.
ACCESSION E27176
VERSION E27176.1 GI:13026394
KEYWORDS JP 1999187879-A/1.
SOURCE Nicotiana paniculata
ORGANISM Nicotiana paniculata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 1950)
Shigehiro, Y. and Toshiyuki, K.
Novel INPS gene derived from nicotiana
Patent: JP 1999187879-A 1 13-JUL-1999;
JAPAN TOBACCO INC
OS Nicotiana paniculata
PN JP 1999187879-A/1
PD 13-JUL-1999
PF 26-DEC-1997 JP 1997359773
PR
PI SHIGEHIRO YAMADA, TOSHIYUKI KOMORI
PC C12N15/09//C12N5/10, (C12N15/09, C12R1:91), C12N15/00, C12N5/00,
PC (C12N15/00, C12R1:91)
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 92..1703.
FEATURES
source Location/Qualifiers
1..1950
/organism="Nicotiana paniculata"
/mol_type="genomic DNA"
/db_xref="taxon:62141"

ORIGIN
Query Match 71.6%; Score 1097.8; DB 6; Length 1950;
Best Local Similarity 82.3%; Pred. No. 8.3e-289;
Matches 1861; Conservative 0; Mismatches 272; Indels 0; Gaps 0;
Qy 1 ATGTCATCGAGAAATTTAAGTAGAGAGTCCTAATGAGTACACCGAGACTGAGATT 60
Db 92 ATGTTATTAAGAAATTTAAGGTGAGACCCCAAGTTAAGTACACCGAAAGTGAATTT 151
Qy 61 CAGTCCGTGTACAACTACGAAACCAACGAACTTTGTTCAGAGAACAGAAATGGACCTAT 120
Db 159 CACTCGTGTATGATTAATCAACCACTGAGTATGTTCAATGAGAGAAATGGACATAT 211
Qy 121 CAGTGAATGTCAAAACCAATCCGTCAATCAATTTAAACCAACCAACCAATGTTCA 180
Db 212 CAATGACCGTCAAGGCTAAGACTGCAATATGAGTTCAAGCTGATGTTATGTTCC 271
Qy 181 AAATGGGGGGTATGTTGTGGGTGGGGTGAACAAAGGCTTACCTTACCCGTGTGT 240
Db 272 AAATTAAGGGGTATGCTTGTGATGGGGTGAACAAATGTTCAACCTTGAACGGTGT 331
Qy 241 GTTATTTGCTAACAGAGAGACATTTTCATGGGCTACAAAGACAAAGATTCACAAAGCAAT 300
Db 332 GTTATTTGCTAACAGAGAGAGATTTTCATGGGCTACAAAGATTAAGTGTGACAAAGCAAT 391
Qy 301 TACTTTGGCTCCCTCAACCAAGGCTCAGTATTCAGATTTGAGTCTTTCAGAGAGAGAA 360
Db 392 TACTTTGGCTCTCTTAATCAAGGCTTCTACTATTCAGATTTGAGTCTTTCAGATTCAGAGAG 451
Qy 361 ATCTATGCCCATTAAGAGTCTGCTTCAATGTGTTAATCTGACGACATTTGTGTTGGG 420

Db 452 ATCTATGCTTCATTTTAAAGCCTCCTTCCAAATGCTCAATTCAGATACCTAGTGTGGA 511
Qy 421 GGAATGGATATACGACAACTGAACCTGCTGATGCCATGCCAAGGCAAAAGTGTTCAC 480
Db 512 GGAATGGATATACGACAACTGAACCTGATGATGATGATGATGATGATGATGATGAT 571
Qy 481 ATGATTTGACAGAGAGTGGGCTTTCATGAAATTCATGATGATGATGATGATGATGAT 540
Db 572 ATGATTTGACAGAGAGTGGGCTTTCATGAAATTCATGATGATGATGATGATGATGAT 631
Qy 541 TAGACCCCGATTTTCAATGCTGCTCCCAACGAGAGAGCGTCCCAACAGCTGATTAAGGC 600
Db 632 TATGACCTCGATTTTCAATGCTGCTCCCAACGAGAGAGCGTCCCAACAGCTGATTAAG 691
Qy 601 ACAAGCAAGAGCAAGTTCAGCAAAATCATCAAGATCATCAAGAGCTTTAAGAGACACC 660
Db 692 ACCAAGAAAGAAACAATTCATCAAAATCATTAAGATTAATTAAGAGAAAGAAC 751
Qy 661 AAAGTGAACAGGTGTTGCTCTGTGACTGCCAACAGAGAGATTAAGATTTGTT 720
Db 752 AAAGTGAACAGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 811
Qy 721 GTAGGCTTAATGACACCATGAGAAATCTTGGCTGCTGTTGAGACGAATGAGGCTGAG 780
Db 812 GTTGAATTAATGACATTAAGAAACCTTTGCTTCTGTGACGAAATGAGCTGAA 871
Qy 781 ATTTCTCTTCCACCTTGTATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 872 ATATCTCTTCCACCTTGTATGATGATGATGATGATGATGATGATGATGATGAT 931
Qy 841 GGAAGGCTTCAAGACCTTTTGTACAGGCTGATGATGATGATGATGATGATGATGAT 900
Db 932 GGAAGGCTTCAAGACCTTTTGTACAGGCTGATGATGATGATGATGATGATGATGAT 991
Qy 901 TTGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 992 TTGATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1051
Qy 961 TTTCTTGGGGGGTGTATCAAGCCCAATCTATATGATTAATCAACCATCTGGGAAC 1020
Db 1052 TTCTTGTGAGAGTGTATTAAGCAACATCATATGATGATCAACCATTTGGGTAAC 1111
Qy 1021 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1112 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1171
Qy 1081 AAGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 1172 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
Qy 1141 GACCATGTTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1232 GACCATGTTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1291
Qy 1201 GAGTACATTCAGAGATTAATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1292 GAGTACATTCAGAGATTAATGATGATGATGATGATGATGATGATGATGATGAT 1351
Qy 1261 GAGATTCCTTTTGAAGTGTCTCTATTAATCTTGAAGTGTGTGTGTGTGTGTGT 1320
Db 1352 GAGATTCCTTTTGAAGTGTCTCTATTAATCTTGAAGTGTGTGTGTGTGTGTGT 1411
Qy 1321 ACTAGATTCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Db 1412 ACCGCAATTCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1471
Qy 1381 ACCATTCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
Db 1472 ACTATTCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1531
Qy 1441 GCAATTCAGAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1500

Db	1532	GGACCTCTCAAAACAGGCGCAATGCTTTGAGAACCAATTTGAGGCGCTTGTTGTGACCTTGCA	1591
Qy	1501	CCAGAGATTAACATGATTTCTCGAGTACCAAGTCA	1533
Db	1592	CCAGAGAACCAACATGATTTCTCGAATTAACAATGA	1624
RESULT 7			
AB009881			
LOCUS			
DEFINITION	AB009881	1959 bp	mRNA linear
ACCESSION	AB009881		PLN 13-UTL-2000
KEYWORDS	AB009881.1	GI:8096265	
SOURCE	myo-inositol 1-phosphate synthase.		
ORGANISM	Nicotiana tabacum		
REFERENCE	Nicotiana tabacum		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
TITLE	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	asterids; Lamiales; Solanales; Solanaceae; Nicotiana.		
	1 (sites)		
	Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.		
	Screening of wound-responsive genes identifies an immediate-early		
	expressed gene encoding a highly charged protein in mechanically		
	wounded tobacco plants		
	Plant Cell Physiol. 41 (6), 684-691 (2000)		
	20399434		
	10945337		
	2 (bases 1 to 1959)		
	Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.		
	Direct Submission		
	Submitted (18-DEC-1997) Kojiro Hara, Nara Institute of Science and		
	Technology, Plant Molecular Breeding; Takayama 8916-5, Ikoma, Nara		
	630-0101, Japan (E-mail:k-hara@bs.aisc-nara.ac.jp,		
	Tel:+81-743-72-5653(ex.5653), Fax:+81-743-72-5659)		
FEATURES			
source	location/Qualifiers		
	1..1959		
	/organism="Nicotiana tabacum"		
	/mol_type="mRNA"		
	/db_xref="taxon:4097"		
	/tissue_type="leaf"		
	94..1626		
	/codon_start=1		
	/product="myo-inositol 1-phosphate synthase"		
	/protein_id="BA095788.1"		
	/db_xref="GI:8096266"		
	/translation="MEINPKVESPNVYKTESIHSVYDYOTTELVEHEKNGTYOMTVF		
	KPKTVKPEKIDVHPKLGVMVGMGNGSTLNGVIANEGISMATKDVQOQANF		
	GSULTASTIRGSRNGEETIAPFKSLPMWNPDDVYREGMISGMNLADAMARKVPE		
	IDLQKVLPRYMESWPLPGIYDPDFIADNQSRRANVTKGKQIDQIADILEFKE		
	KNKRVKLVWNTANTRSNVNVGANDMENLFSVDNENAEISPTLYAIACTLENNV		
	PIINSQNTPEVGLIDLAIKRNTLIGDDFSGGTGKSKVLDPLVAGIKPISVSG		
	VNHLNNDGQMLNAPOTRSKEISKSNVDDVSNALTYRGEHPDHVVITKVPYVA		
	GDSKRAMPEYSEIFMGKNTVIANTEQDSLIAPILDIIVLAELSTRITLAKGEE		
	GKFSHFHPVATITLSTYTKAPLVPPEPTPVNALS KORAMLENTILRACGLAPENNMTLE		
	YK"		
	1959		
	/note="15 a nucleotides"		
polya_site			
ORIGIN			
	Query Match	71.3%;	Score 1093; DB 8; Length 1959;
	Best Local Similarity	82.1%;	Pred. No.1.7e-287;
	Matches 1258; Conservative	0;	Mismatches 275; Indels 0; Gaps 0;
Qy	1	AGTTCATCGAAGATTTTAAGTGAAGAGTCTCTAATGTGAAGTACACCGAGACGTGAAGTT	60
Db	94	ATGTTTATTGAAGACTTTAAGGTGAGAGCCCAACGTTAAGTACACCGAAAGTGAAGTT	153
Qy	61	CAGTCGCGTGTACAGTACAGAAACCAACCGAACTTGTTCACAGAAACAGAAATGGACCTAT	120
Db	154	CAGCTGCTGCTATGATTTATCAACCACTGAGTTAGTTCAATGAGAGAGAAAATGGAGCTTAC	213
Qy	121	CAGTGAATTTGTCAAAACCAATCGGTGCACTACCAATTTAAACCAACCAACCATGTTTCA	180

Db	214	CAATGACATGCTCAAGCCTTAAGACTGTCAAACTGATGTTCAATGTTCC	273
Qy	181	AAATTTGGGGGTGATGCTTGTGGGTTGGGGGTGAAAACAACGGCTCTACCTCACCAGGTGT	240
Db	274	AAATTTAGGGGTTATGCTTGTGTGGATGGGGGTGAAAACAATGGTTCAACCTTGACCGGTGTGT	333
Qy	241	GTTTATTCGTAAACAAGAGACACTTTCATGAGGCGTCAAAAGACAAGATTCAACAAGCCAAAT	300
Db	334	GTTATTTGTAAACAAGAAAGAAATTTCAATGGGCCCAAAAGATAAGGTGCAACAAGCCAAAT	393
Qy	301	TACTTTGGGCTCCCTCACCCAAAGCCTCAGACTATTTGAGTTGGATCTTTCAGAGGAGAGAA	360
Db	394	TACTTTGGGCTCTTACTCAGGCTTCTACTTTTGAATTTGAGTTGGGCTTTCATAGGAAGAG	453
Qy	361	ATCTATGCCCCATTCAAGAGTCTGCTTCCATGAGTTAATCTGACGACATTTGTGTTGG	420
Db	454	ATCTATGCTTCATTTAAAGGCTCTCTTCCATGAGTCAATCCAGATGACGTATGTTTGA	513
Qy	421	GGATGGGATATCAGCAACATGAACTCGGTGATGCTCATGGCCAGGCGCAAGAGTGTGAC	480
Db	514	GGATGGGACATCACCGGACATGAATTTGGCAGATCCCATGGCAAGGGCTTAAGATTTTGAT	573
Qy	481	ATCGATTTGCAAGAGCATGTTGAGGCGTTTACATGGAAATCAATGTTCCACTCGCCGGATTC	540
Db	574	ATTGATCTCAAAAAGCAAGTTGAGGCCCTCACAATGAAATCTATGTTCCACTACCTGGTATTC	633
Qy	541	TACGACCCGGATTTTCATTTGCTGCCCAACAAGAGAGCGTGCCAACAACGTGATTTAAGGC	600
Db	634	TATATCTCTGATTTTCATTTGCTGCTAACCAAGAGGTCAAGTGCACAACGTATCAAGA	693
Qy	601	ACAAAGCACAAGCAAGTTACAGCAAAATCATCAAGAATCATCAAGGCGTTTAAAGAACCAAC	660
Db	694	ACCAAGAAAGAAACAATTTGATCAAAATCATTAAGATATTAGGGAGTTTAAAGAAAGAAC	753
Qy	661	AAAGTGGACAAGGTGGTTGCTGTGGACGTGCCAACAAGAGAGGTATAGCAATTTGGTT	720
Db	754	AAAGTGGACAAGGTGGTATTTTGTGGACTGTCTTACACTGAAGAATACAGTAAATGTGTT	813
Qy	721	GTAGGCTCTTAAATGACACCATGAGAAATCTCTTGGCTGTGTGACAGAAATGAGGCTGAG	780
Db	814	GTTGGACTTAAATGATACATAGGAATAATCTTTTGCTGTGTGACAGAAATGAGGCTGAA	873
Qy	781	ATTTCTCTCTCACACTTGTATGCCATTTGCTGTGTGATGAAAATGTTCCTTTCAATTAAT	840
Db	874	ATATCTCTCTTCCACTTTGTATGCTATTTGCCGTGCAATTTTGAAAATGTGCCCTTTATTAAT	933
Qy	841	GGAAAGCCCTCGAACAATTTTGTATCCAGGGGTGATTAATCTTGCCATCGCAGGAACACT	900
Db	934	GGAAAGCCCAAGAACACCTTTGTGTCCAGGCTCAATGATTTGGCCATTCAGAGAAACACA	993
Qy	901	TTGATTTGGTGAAGATGACTTCAGAAAGTGTGACAGCAAAATGAATCTGTGTTGGTTGAT	960
Db	994	CTGATTTGGTGTATGACTTTTAAAGATGTGTAAACCAAAATGAAGTCAGTGCTGTTGAT	1053
Qy	961	TTTCTTTGGGGGCTGTGATCAAGCCCAACATCTATAGTTAGTCAACCATCTGGGAAAC	1020
Db	1054	TTCTTTGTGGAGCTGTGATTAACCAACATCAATTTGAGCTTCAACCATTTGGGTAAAC	1113
Qy	1021	AATGATGATGAATCTCTCGGCTTCCAAACCTTCGCTCCAGGAATGTTCCAAAGAC	1080
Db	1114	AATATGGAATGAATCTGTCTGCCCTTCAACCTTTCGCTCAAAAGGAGATCTCAAAAAGT	1173
Qy	1081	AAAGTTGTGACATGATGTCMAACAGCAATGCCATCTTATGAGCTGTGGAACATCCC	1140
Db	1174	AATGTTGTGATGACATGTGTTCAAGCAATGCCATCTTATGAGCTGTGAGAGACACCT	1233
Qy	1141	GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGGGATAGCAAGAGACCATGGAT	1200
Db	1234	GACCATGTTGTTGATTAAGTATGTGCTATATGTGGGAAGACCAAAAGGGCAATGGAT	1293
Qy	1201	GAGTACACTTCAAGATATTCATGGGTGGAAAGAACACATTTGTTTTCACAAACATGT	1260

Db 1294 GAGTACACATCTGAGATTTTCATGAGGAGAAACACATGTTTTCACAAATCTGT 1353
Qy 1261 GAGATTCCTTTTACTGCTGCTATTAATCTTGAGACTGGTCTTCTGTGAGCTGAGC 1320
Db 1354 GAGGATTCCTTTTGGCTGCTGCTCAATTAATCTTGATTTGGTCTTCTGTGAGACTGAGT 1413
Qy 1321 ACTAGATTCAGATTTAAAGCTGAAATGAGGAGAAATTCACATCTTCCACCCAGTTGCT 1380
Db 1414 ACCCGATTCAGCTCAAGAGCTGAGAGAGAGGTAAGTTCACATCTTCCACCCCGTGGCT 1473
Qy 1381 ACCATTCAGCTATCTGACCAAGGCTCTCTGTGTTCCACCGGTAACCAAGTGTGAAT 1440
Db 1474 ACTATCTCAGCTACCTTACCAAGGCTCTCTGTGTTCCACCAAGTGTGAAT 1533
Qy 1441 GCATTCGAAAGAGGCTCAATGCTGGAATACATATAGAGGCTTGTGTGATTTGGCC 1500
Db 1534 GCACTCTCAAGAGAGGCGAATGCTTGAGAACATATAGAGGCTGTGTGATTTGCA 1593
Qy 1501 CCAGAGATTAAGATGATTTCTCGAGTACAAAGTGA 1533
Db 1594 CCAGAGAACACATGATTTCTGGAATACAAATGA 1626
RESULT 8
BT013759 1954 bp mRNA linear PLN 11-MAY-2004
LOCUS Lycopersicon esculentum clone 132642F, mRNA sequence.
DEFINITION BT013759
ACCESSION BT013759.1 GI:47105174
VERSION FLI_CDNA.
KEYWORDS Lycopersicon esculentum (tomato)
SOURCE Lycopersicon esculentum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1954)
Kirkness, B.F., Wang, W. and Vazelle, A.
REFERENCE Direct Submission
AUTHORS Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
TITLE Medical Center Drive, Rockville, MD 20850, USA
JOURNAL Location/Qualifiers
FEATURES
Source 1..1954
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="132642F"
/issue_type="Final"
/note="TZXB41"

ORIGIN

Query Match 69.7%; Score 1069; DB 8; Length 1954;
Best Local Similarity 81.1%; Pred. No. 6.3e-281;
Matches 1243; Conservative 0; Mismatches 290; Indels 0; Gaps 0;
Qy 1 ATGTCATCGAGATTTTAAGTAGAGAGTCTAATGTGAATAGACCGAGCTGAGAT 60
Db 107 ATGTTATTTAGAACCTTTAAGGTTAGAGCCCAATGTTAAGTACACAGAGTAGAAT 166
Qy 61 CAGTCGCTGTAACAATAAGCAACCAAGCTTGTTCAGAGACAGAGATGGCACTAT 120
Db 167 CACTCTGTGTATGATTAATCAACCACTAGTGTTCATAGAGAAAAAATGGGACATTT 226
Qy 121 CAGTGGATTTGCAAAACCAATCCGTCACATCAATTTAAACCAACACCATTTTCA 180
Db 227 CAATGCACTGTAAACCAAAACCTGTCAAAATGTGAATCAAAACCTGATGTCATGTC 286
Qy 181 AAATGGGGGTGATCTGTGGGTTGGGGTGAACAACGCTCAACCTCAACCGGTGT 240
Db 287 AAATTTAGAGATTAATCTGTGGATGGGTTGAAACCAATGTTCAGCTTGAACCGAGGT 346
Qy 241 GTTATTTGCTAAACAGAGACATTTCAATGGGCTACAAAGACAAATTTCAACCAAT 300
Db 347 GTTATTTGCTAAATAGAGAAATTTCAATGGGCTACAAAGATTAAGTGCAGAACCAAT 406

Qy 301 TACTTGGCTCCCTCAACCCAGAGCTGAGCTATTCAGATTGATTCCTTCCAGGAGAGAA 360
Db 407 TACTTGGCTCCCTTACACAGAGCTCTACTATTCAGATTGAGGCTTTTCAATGAGAGAG 466
Qy 361 ATCTATGCCCAATTCAGAGTCTGCTTCCATAGTTATCTCTGACAGACATATGTTGGG 420
Db 467 ATCTATGCTCCCTTTAAGACATCTTCCATAGTCAATCCAGATATATGATTTTGA 526
Qy 421 GATGGGATATACAGCAATGAACCTGAGTATGCCATGAGGCAAGGCAAGGTTTGA 480
Db 527 GATGGGACATACAGCAATGAATTTGGAGATGACATGAGCAAGAGCTAAAGTTTGA 586
Qy 481 ATGATTTGCAAGAGAGTGAAGCTTACATAGAAATTCATGTTTCACTCCCGAATC 540
Db 587 ATGATCTTCAATGAATGAACCTGAGGCGTACATGAATCATGATCTCCACTGCTGATC 646
Qy 541 TACGACCCGGAATTTCTCTGCTCCCAACAGAGAGGCTGCCAACAACGTATTAAGGC 600
Db 647 TATGACCTGATTTCTATCTGCTCCCAACAGGCTCAGTGTACATATCATCAAGGA 706
Qy 601 ACAAGCAAGAGCAAGTTGACGAAATCATCAAGACATCAAGCGTTTAAGAGACACC 660
Db 707 ACCAAGAAAGAACAAATGATCAAAATTAATTAAGATATTAGGAGTTTAAGGAGAACAC 766
Qy 661 AAAGTGAACAGAGTGTCTCTGTGAGTCCCAACAGAGAGTATAGCAATTTGGTT 720
Db 767 AAAGTGAACAGAGTGTCTCTGTGAGTCCCAACAGAGTATAGCAAGTGTGGCT 826
Qy 721 GTAGGCTTAATGACCAATGAGAAATCTCTGGCTGCTGTGAGCAAGAAATGAGCTAG 780
Db 827 GTTGGCTTAATGATACATGAGAAACCTTTGCTCTGTGATGAATGAGTGA 886
Qy 781 ATTTCTCTTCCACCTTGATGATCCATGCTGTGATGAGAAATGTTCTTTCAATAT 840
Db 887 ATATCTCTTCCACCTTGATGATCTATGCTGTATTTCTGAAATATGACCTTCAATAT 946
Qy 841 GGAAGCCTCAGAACACTTTGTACAGAGGCTGATTTGATCTTCCATCGAGAACACT 900
Db 947 GGAAGCCTCAGAACAAATCTTTGTCCAGGCTCATTTGATTTGGCATCAAGAGAACACT 1006
Qy 901 TTGATTTGGTGAAGTACTCTCAAGTGTGACCAAAATGAATCTGTGTTGAT 960
Db 1007 TTGATTTGGTGAAGTACTTTAAGTGTGACCAAAATGAATGATGATGCTGTGAT 1066
Qy 961 TTTCTTTGGGGCTGTATCAAGCAACATCTATAGTTAGTAAACCATCTGGGAAAC 1020
Db 1067 TTCTTTGGAGCTGTATTAAGCAACATCAATCGAGCTACACCATTTGGGTAAAC 1126
Qy 1021 AATGATGTATGATCTCTGGCTCCACAAACCTTGGCTCCAGAGAAATCTCAAGAC 1080
Db 1127 AATGATGTATGATCTCTGGCTCCACAAACCTTGGCTCCAGAGAAATCTCAAGAC 1186
Qy 1081 AAGCTTTGAGAGATGATGTAACGCAATGCAATCTTATGAGACCTGTGAGAGATCT 1140
Db 1187 AAGCTTTGAGAGATGATGTAACGCAATCTTATGAGACCTGTGAGAGATCT 1246
Qy 1141 GACCATGTTGTTATTAATGATGCTTACATGAGGAGATGAGCAAGACCATGAT 1200
Db 1247 GACCATGTTGTTGATTAATGATGCTTACATGAGGAGATGAGCAAGACCATGAT 1306
Qy 1201 GAGTACATTAAGATATTAAGGAGGAGAAACACATTTGTTTTCACAAACATGT 1260
Db 1307 GAGTACATTAAGATATTAAGGAGGAGAAACACATTTGTTTTCACAAACATGT 1366
Qy 1261 GAGATTCCTTTTAAAGCTCTTATTAATCTTGAATCTTGTCTTGTGAGTGAAC 1320
Db 1367 GAGATTCCTTTTAAAGCTCTTATTAATCTTGAATCTTGTCTTGTGAGTGAAC 1426
Qy 1321 ACTAGATTCAGATTTAAAGCTGAAATGAGGAGAAATTCACATCTTCCACCCAGTTGCT 1380
Db 1427 ACCCGATTCAGCTCAAGGCTGAGAGAGGATTAAGTTCCACTCTTCAATCCGTGGC 1486

QY 1381 ACCATTCTCAGCTATCTGACCAAGAGCTCCTGTGTTCCACCGGTTACACCAAGTGTGAT 1440
| | | | |
DB 1487 ACCATCTCAAGTATCTTACCAAGAGCTCCTTTGTATACCAACGAGTACACAGTGTGAT 1546
| | | | |
QY 1441 GCATTGTCAAAAGCAGCGTGCATGCTGAAAACATPATGAGGGCTTTGTGTTGATGGCC 1500
| | | | |
DB 1547 GCCCTTTCAAAAGCAGAGGCAATGCTCGTGAACATATTTGAGGCTTGTGTGACTGGCA 1606
| | | | |
QY 1501 CCAGAGATTAACATGATTTCTGAGATCAAGTGA 1533
| | | | |
DB 1607 CCAGAGAACACATGATTTGGAATACAAATGA 1639
| | | | |
RESULT 9
BT013505 1959 bp mRNA linear PLN 11-MAY-2004
LOCUS BT013505
DEFINITION Lycopersicon esculentum clone 132193f, mRNA sequence.
ACCESSION BT013505
VERSION BT013505.1 GI:47104920
KEYWORDS FLI CDNA.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1959)
Kirkness, E.F., Wang, W. and Vazelle, A.
REFERENCE Direct Submission
Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Journal Medical Center Drive, Rockville, MD 20850, USA
FEATURES
source
1..1959
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="132193f"
/issue_type="frule"
/note="TXBFS1"
ORIGIN
Query Match 69.2%; Score 1061; DB 8; Length 1959;
Best Local Similarity 80.8%; Pred. No. 9,7e-379;
Matches 1238; Conservative 0; Mismatches 295; Indels 0; Gaps 0;
QY 1 AGTTCATGAGAGATTTTAAGTAGAGAGTCCTAATGTGAAGTACACGAGCTGAGATT 60
| | | | |
DB 77 AGGTTATTAAGAAATTTTAAGGTGAAGGCCAAATGTGAATTTGAGATGAATTT 136
| | | | |
QY 61 CAGTCGCTGTAACAACCTACGAACCAACGAACTTGTTCAGAGAAACAGAAATGGCACCTAT 120
| | | | |
DB 137 CATTCTGTATGATGATATGAACCAAGAGCTTGTTCATGAAGAGAAATGGAATCTTAT 196
| | | | |
QY 121 CAGTGAATGTCAAAACCAATCCGTCAACTACCAATTTTAAACCAACACCTATTTCCA 180
| | | | |
DB 197 CAATGATTTGTTAAGCTTAACCTCAATATGAATTTAAACCTATACCCATGTGCA 256
| | | | |
QY 181 AATTTGGGGGTGAGTGTGGGTGGGGTGAACAACGCTCTACCCCTACCGGTGAT 240
| | | | |
DB 257 AATTTGGGGGTATGCTTGTGGATGGAGGAAACATGTCTCAACATTTGACTGAGGT 316
| | | | |
QY 241 GTTATTTGCTAAACAGAGACATTTATGAGGCTTACAAAGACAAAGATTCAACAGCAAT 300
| | | | |
DB 317 GTTATTTGCGAATCGAAGAGAAATTTATGAGGCAACGAAAGAAAGTGAACAGCAAT 376
| | | | |
QY 301 TACTTTGGCTCCTCAACCCAGCTCAGCTATTTCAGTTGATTCCTTCCAGGAGAGAA 360
| | | | |
DB 377 TATTTGGGTCTCTTACTCAGGACATCAACCATTCGAGTGGGTCTTTCAATGCGAAG 436
| | | | |
QY 361 ATCTATGCCCCATTCGAAGAGTCTGCTTCCATGTGTTAATCTGACGACATGTTGGG 420
| | | | |
DB 437 ATCTATGACCCCTTCAAAAGCTCTTCCATGTGTCAACCAACGATGTAGATTGGA 496
| | | | |
QY 421 GATGGATATTCAGAACTGAACCTGCTGATGCCATGGCCAGAGGCAAGGTGTTGAC 480
| | | | |

DB 497 GGATGGGACATTAGACATATTAATTTGGCAGATGCTATGGTCAGGGCTTAAGGTTTCA 556
| | | | |
QY 481 ATTCATTTGCAAGACAGCTTGAAGGCTTACATGAAATCCATGTTTCCACTCCCGGAATC 540
| | | | |
DB 557 GTTGAATCTGCAAAAGCACTGAGGCCCTTACATGAAATCAATGTTCCCTTCTCGTATC 616
| | | | |
QY 541 TACGACCCGATTTTCAATGCTGCTCCAAACAGAGAGCGTCCCAACACGTATTAAGGCC 600
| | | | |
DB 617 TATGACCCGACTTATGACGAGCTTAAACAGAGACGCTGCCAACACGTGATCAAAAG 676
| | | | |
QY 601 ACAAGCAAGACCAAGTTTCAAGCAATCATCAAGCATCAAGCGCTTAAAGAACCCACC 660
| | | | |
DB 677 ACCAAGAAAGAACAGTTGAACCAATTTGTAAAGATTTAGGAGTTCAAGAGAAAGAAC 736
| | | | |
QY 661 AAAGTGAACAAGTGGTGTCTGCTGTGACCTGCCAACAGAGAGGTATAGCAATTTGTT 720
| | | | |
DB 737 AAGTGAACAGATATGAGTGTCTTATGACCTGCCAACACGAAAGTACATGATGTGTT 796
| | | | |
QY 721 GTAGGCTTTAATGACACCATGAGAAATCTTTGCTGTGTGACAGAAATGAGCGCTGAG 780
| | | | |
DB 797 GTTGGCTTTAATGACACCATGAGAAACCTTTTAAAGTGTGCTGTGATGAATGAGCGTGA 856
| | | | |
QY 781 ATTTCTCTTCACTTGTATGATGCCATTTGCTGTGATGAAATGTTCTTCAATAT 840
| | | | |
DB 857 ATATCTCTTCACTGATGCTATTTGCTGTGATGAAATGCTTTCATCAAC 916
| | | | |
QY 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCATGCGAGAACT 900
| | | | |
DB 917 GGAAGCCCTCAGAACACTTTTGTTCAGAGTCTTATTTGATTTGGCATTAAGGAACACT 976
| | | | |
QY 901 TTGATTTGTGAGATGACTCTCAAGAGTGTGACACCAAAATGAATCTGTGTTGAT 960
| | | | |
DB 977 TTAATTTGGTGTGATGACTTTTAAAGTGTGCAACCAAGATGAATGCGTGTGTTGAT 1036
| | | | |
QY 961 TTTCTTTGGGGGCTGTGATCAAGCAACATCTAATAGTTAATCAACCATCTGGGAAC 1020
| | | | |
DB 1037 TTCTTTGTGAGCCGCTATTAAGCAACGTCAATAGTGAATCAATCACTTGGGTAAAC 1096
| | | | |
QY 1021 AATGATGATATGATCTCGGCTCCACAAACCTTCCGCTCCAGAAATCTCAAGAGC 1080
| | | | |
DB 1097 AATGATGATATGATCTTCTGCTCTCAACCTTCCGCTCAAGAGATCTCAAAAGT 1156
| | | | |
QY 1081 AAGTGTGACGATATGCTCAACAGCAATGCAATCTCTATAGCTGTGTAACATCCC 1140
| | | | |
DB 1157 AATGTTGTGATGACATGTTGCTAGCAACGCAATCTCTATGAATCTGGCGACACCT 1216
| | | | |
QY 1141 GACCATGTTGTTATTTAATATGCTCTTACATAGGGATATGCAAGAGCCCATGAT 1200
| | | | |
DB 1217 GACCATGTTGCTGATCAAGATGTTCCATATGTTGAGACAGCAAGAGGCAATGAT 1276
| | | | |
QY 1201 GAGTACACTTCAAGATATTTCAAGGGTGAAGAAACACATTTGTTGCAACACATGT 1260
| | | | |
DB 1277 GAGTACATCTCAGATTTTCAATGCGGGAAGAGACATATGATGCAACACTTGT 1336
| | | | |
QY 1261 GAGATTCCTTTTAACTGCTCTTATTTCTTGAATGCTGCTTCTGTGACTGAGC 1320
| | | | |
DB 1337 GAGGACTCTCTTTGGCACTCCAAATATCTTGGATTTTGCTCTCTCTCAACCTCAGC 1396
| | | | |
QY 1321 ACTGAATTCAGTTTAACTGAAGTAAATGAGGAAATTCACATCTCAACCAAGTGTCT 1380
| | | | |
DB 1397 ACTGCAATTCACCTCAAGCTAAAGGAGGAGGAGGAGTTCACCTCTTCCACCTGTGGCT 1456
| | | | |
QY 1381 ACCATTTCCAGCTATCTGACCAAGGCTCTGTGTTTCCACCGGTTACACAGTGTGAT 1440
| | | | |
DB 1457 ACTATTTCAAGCTTACTTCAAGAGCTCCCTGTGATCAACAGTACTCAAGTGTGAT 1516
| | | | |
QY 1441 GCATTGTCAAAAGCAGCGTGCATGCTGAAAACATPATGAGGGCTTTGTGATTGGCC 1500
| | | | |
DB 1517 GCCCTTTCAAAAGAGGCAATGCTTGAATATPATGAGCTTGTGTGTTGGCA 1576
| | | | |
QY 1501 CCAGAGATTAACATGATTTCTGAGTACAAAGTGA 1533
| | | | |

Db 1577 CCAGAGAACATGATTTGGATACAAATGA 1609

RESULT 10
AF284065 1845 bp mRNA linear PLN 30-JUL-2003
LOCUS Sesamum indicum myo-inositol 1-phosphate synthase mRNA, complete cds.
DEFINITION AF284065.1 GI:9858815
VERSION AF284065.1
KEYWORDS
SOURCE Sesamum indicum (sesame)
ORGANISM Sesamum indicum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; Lamiales; Pedaliaceae; Sesum.
REFERENCE 1 (bases 1 to 1845)
Chun, J. A., Jin, U. H., Lee, J. W., Yi, Y. B., Hyung, N. I., Kang, M. H., Pyee, J. H., Suh, M. C., Kang, C. W., Seo, H. Y., Lee, S. W. and Chung, C. H.
Isolation and characterization of a myo-inositol 1-phosphate synthase cDNA from developing sesame (Sesamum indicum L.) seeds: functional and differential expression, and salt-induced transcription during germination
Plantia 216 (5), 874-880 (2003)
JOURNAL MEDLINE 22511750
PUBMED 12624775
REFERENCE 2 (bases 1 to 1845)
Jin, U. H. and Chung, C. H.
AUTHORS Direct Submission
TITLE Submitted (01-JUL-2000) Division of Biotechnology, Faculty of Life Science and Resources, Dong-A University, Ha-Dan-Dong, Sa-Gu-Gu, Pusan City, Pusan 604-714, South Korea
JOURNAL Location/Qualifiers
FEATURES
source
1.1845
/organism="Sesamum indicum"
/mol_type="mRNA"
/db_xref="taxon:4182"
/feature_type="sequence"
/note="authority: Sesamum indicum L."
51.1583
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AA601148.1"
/db_xref="GI:9858816"
/translation="MFIESFKVESPNVKTGEIHSVNYETTELVESRNCYQIVKRTVYERKIDTHVPLKGVMLVGMGNGSTLTGVINRISISWATDXTQOANVFGSLTOASSIRVGSFNGEIEIYAPFKSLPWNPDVYFGMDISNNMLADAMRAKVLIDLOKOLRYMEHMYPLPGIYDPDPIAANGSRANNVIGTKKEOVQOIIKMRDPEOKRDKVYLTANTERYSNVVGLNDTPESLMAVSERBEAISPTETAIACVPEVNPINGSPOUTFVPGILDIAIQNSLIIGDVFSGQTKMSVLVDPLVAGIPEPTIVSYHLGNNDDMNLSAPQTFRSKSIKSNVDDVNASNGILYERBEHFDHIVIKYEVYGSKRAMDEYTSIEIFWGSSTIVLHNTCCDSILAAPIIIDLVLAEISTRIDKKEGEGFHSFHPVATILSYLTAKAPLVPGTPVNVNALSQRAMENILRACVGLAPENNMLEYK"

ORIGIN

Query Match 68.2%; Score 1045; DB 8; Length 1845;
Best Local Similarity 80.1%; Pred. No. 2,3e-274;
Matches 1228; Conservative 0; Mismatches 305; Indels 0; Gaps 0;

QY 1 AAGTCATGAGAAATTTAAGTAGAGAGTCCCTAATGTGAATGACCGAGACTGAGATT 60
DB 51 AAGTTCATGAGAGCTTCAAGGTTGAGACCCCAATGTCAAGTACCGAGGGGTGAGATT 110
QY 61 CAGTCGCTGATCAACATGAGAAACCAACGAACTGTTACGAGAAACGAGATGGACCTAT 120
DB 111 CACTCGTCTACACACTACGAGACTACAGAGCTTGTTCAATGAGAGCAGAAACGCACTAC 170
QY 121 CAGTGCATTTGTAACATCCCAATCCGTCATCACTACCAATTTAAACCAACCAACCTGTTCA 180
DB 171 CAATGATATTTGTCAAGCCCAAGACTGTCAATACGAGTTCAAACTGATCTACATGTCCTCC 230
QY 181 AATTTGGGGGTATGCTTTGGGTTGGGGTGGAAACAGCGCTCTACCTCAACCGGTGCT 240

DB 231 AATTTAGGGGTATGCTTTTGGATGCGGAGGAACAATGTTCACTCACTGCGGT 290
QY 241 GTTATTTCTAACAGAGAGACATTTTCATGCGCTTACAAAGACAGATTTCAACAGCCAAAT 300
DB 291 GTTCATTTCAAAAGAGGAGATTTTCATGCGGCAACAAGATTAAGTGCACAGCCCAAT 350
QY 301 TACTTTGGCTCCCTCAACCCACCTCAGCTATTTCAGATTGATTCCTTCCAGGAGAGGAA 360
DB 351 TATTTTGGGTCACTGACACAGAGCATCTCAATCAGATTGTTCTTCAATGAGAGAGAG 410
QY 361 ATCTATGCCCCCAATTCAGAGTCTGCTTCCATGTTTAACTGACAGACATTTGTTGGG 420
DB 411 ATCTATGCCCCCTTCAAAAGCTTCTTCCCATGCTACACCGAGATACCTAGTGTGGG 470
QY 421 GGAATGGATATACGACATGAACTGCTGATGCCATGCGCAAGGCAAAAGGTGTTGAC 480
DB 471 GGATGGGACATCAGCAATATGAAATTTGCTGATGCCATGGCAGGGCAAGGTGTAGAT 530
QY 481 ATCGATTTGACAGAGCATGTTGAGGCTTTCATGGAATCCATGTTCACTCCCGGAATC 540
DB 531 ATTGATCTTCAGAGAGCATGCTCAGGCTTCATGAGAGCATATGCTCCACTCCCTGGAATC 590
QY 541 TACGACCCGGATTTTCACTTGTGCTGCAACCAAGAGAGCGGCCAACCAACGTGATTAAGGC 600
DB 591 TAGATCTCGATTTTCACTAGCCGCAATCAGGATCAGTGTCCCAACAGCTGATTAAGGA 650
QY 601 ACAAGCAAGACAGCATGTCAGCAAAATCATCAAGATCATCAAGCGTTTAAAGAACCC 660
DB 651 ACCAAGAAAGAACAGATTCAACAGATCATCAAGATCATGAGGACCTTTAAGAGCAAAAT 710
QY 661 AAGATGACAAAGTGTGTTCTCTGTGACTGCCACACAGAGAGTATGCAATTTGGTT 720
DB 711 AAGTGGACAAAGTGTGATGCTGTGACAGCAACACAGAAAGATACGTATATGTTGTC 770
QY 721 GTAGGCTTATGACACCATGAGATCTCTTGGCTGCTGTGACAGAAATGAGCTGAG 780
DB 771 GTTGGCTTAAAGATACAGATACAGCCAAAGCTGATGCTCGGTAGAGAGATAGCCGAG 830
QY 781 ATTTCTCTTCCACTTGTATGATCCATGCTGCTGTGATGGAATATGTTCTTCAATTAAT 840
DB 831 AATTCCTCTTCAACCTTGTATGATCCATAGCTGTGTTTTGAAGATGTTCTGTTCAATTAAT 890
QY 841 GGAAGCCTCAGACACTTTTGTACAGAGGCTGATTTGATCTTGCATGCGAGAAACCT 900
DB 891 GCAGGCCCAAAATATCTTTTCCAGGCTGATTTGATTTGGCGATTCAGAGGAAACCT 950
QY 901 TTGATTTGGTGAATGATCTTCAAGAGTGTGACCAACAAATGAATCTGTGTTGTTGAT 960
DB 951 TTGATTTGGTGAATGATTTCAAGAGCGGTCAACCAAGATGAATCAAGTGTAGTTGAT 1010
QY 961 TTTCTTGGGGGCTGTATCAAGCCAACTATATAGTTAGTACCAACCATCTGGGAAAC 1020
DB 1011 TTCTTTGTTGAGCTGTGATCAAGCCAACTGATTTGTAGAGCTATGATCTGGGAAAC 1070
QY 1081 AAGCTTTGAGATATGATGTCACAGCAATGCAATCTCTATAGAGCTGTGTAACATCCC 1140
DB 1131 AATGTGTGAGACATGTTGTCCACAGACGGCATCTTATAGAGCTGTGTAAACCCCC 1190
QY 1141 GACCATGTTGTTGTTATTAAGTATGCTTACGTAGGGGATAGACAGACCATGTGAT 1200
DB 1191 GACCATATTTGTTGTCATCAAGATATGATCGTATGTGGGGGACAGCAAGAGACATATGAGAC 1250
QY 1201 GAGTACATTTGAGATATTTACATGAGTGAAGAAACACATTTGTTTGCACAAACATGT 1260
DB 1251 GAGTACATTTGAGATATTTACATGAGTGAAGAAACATATGTTTGTGCAACACATGT 1310
QY 1261 GAGATTCCTCTTGTGCTGCTCTATTTATCTTGAAGCTTGTGCTTGTGAGCTGAGC 1320

Db 1311 GAGGACTCTCTTGTGGCTGCTCCATCATCTCGATTGGTCCCTAGCTGACTGACG 1370
Qy 1321 ACTAGATCCAGTTTAAGCTGAATGAGGAAAATTCCTCACTCAATCCACCCAGTTGCT 1380
Db 1371 ACTGTAATTCAGCTCAAGCCGAGGAGGAGGCAATTCATTCTGTTCCACCCGGTGGCA 1430
Qy 1381 ACCATTCTCAGCTATCTGACCAAGGCTCCTGTTCCACCGGGTACACCACTGGTGAAT 1440
Db 1431 ACCATCTCAGCTACTCAACCAAGCCCTCTTTGTAACCTCCAGGCACACCGGTGCTCAAC 1490
Qy 1441 GCATTGTCAAGACGCGTGCATATGCTGAAAACATATAGAGGGCTTTGTGTGGATTGGCC 1500
Db 1491 GCACGTCTTAAGCAGCGTCCCTGCTCGAACAATCTTGAAGGGCTTGTGTGGATTGGCT 1550
Qy 1501 CCAGAGATTAACATGATTTCTCGAGTACCAAGTGA 1533
Db 1551 CCAGAAAACATATGATTTTGGAAATACAGTGA 1583

RESULT 11
CPINOIG 1978 bp DNA linear PLN 26-JUN-1995
LOCUS C.paradisi (Macf) INO1 gene.
DEFINITION Z32632.1 GI:602564
ACCESSION Z32632.1 GI:602564
VERSION INO1 gene.
KEYWORDS Citrus x paradisi
SOURCE Citrus x paradisi
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Sapindales; Rutaceae; Citrus.

REFERENCE 1 (bases 1 to 1978)
AUTHORS Holland,D.
TITLE Direct Submission
JOURNAL Submitted (08-APR-1994) Doron Holland, Fruit Tree Breeding and Genetics, Volcani Research, Organization Center, Bet Dagan, 50250, Israel

FEATURES
source Location/Qualifiers
1..1978
/organism="Citrus x paradisi"
/mol_type="genomic DNA"
/db_xref="taxon:37656"
/clone="C38"
/issue_type="fully expanded leaves"
/dev_stage="juvenile plants"
188..1711
/gene="INO1"
188..1711
/gene="INO1"
/protein_id="CAA83565.1"
/codon_start=1
/protein_id="GI:602565"
/db_xref="GI:602565"
/db_xref="GOA:P42802"
/translation="MEIENFKVESPVKXTDHIHSVVDYETTELVHENRNGTYOMIV
KERKYVERKTDVHVYKLGVMVVGWGNSTDLTGVLINREGICMATDQTVQOANYF
GSLTQASAIRVGSYNGEYIAPFKSILPMVNSPDIVYFGWMDISDMMLADAMARAFED
IDLOKOLRPMYSNVPRLPGIYDPRDPIANOGSRANNVIKCTKKOMLOIIXIDIREKE
KAKRDRVVLVLTANTERYSNVYGLNDYTESLASIDKKEAITSSTLYAIXCVLENT
PFINSPONTFVPGGLDLAIRNCLIGSDPKSGQTKMSVLVDPLVGGIKPTISVS
VNHLSNDDMNLSAPQTFPSKESIKSNVVDVWSSNVFPMGLVNRPRMVKVPPYVAI
BRAMDEYTSIEFMGKSTIVLNTCEDSLAAPILIDVLVLAELSTRIQLKKEGSGKF
HSFHPVATILSLTKAPLIVPGRPVVNALSKQRAMLENTLRACVGLAPENNMILBYK"

ORIGIN

Query Match 68.1%; Score 1044.6; DB 8; Length 1978;
Best Local Similarity 81.9%; Pred. No. 3e-274;
Matches 1255; Conservative 0; Mismatches 269; Indels 9; Gaps 4;

Qy 1 ATGTCATGAGATTTTAAGTAGAGAGTCCTAATGAGTACCGAGACTGAGATT 60
Db 188 ATGTTATGAAAATTTCAAGGTGAGACCCCAATGTGAAGTACAGATCATGAGATT 247

Qy 61 CAGTCGTGTACAACTACGAAACCAACCGAATTGTTACGAGAACAGAAATGGCACTTAT 120
Db 248 CACTCTGTGTGTCGATTATGAAACCACTGAGCTTGTTCATGTAGAACAGAAATGGCACTTAT 307
Qy 121 CAGTGGATTTGCAAAACCCCAATCCGTCAACTACCAATTTAAACCAACCACTGTTTCCA 180
Db 308 CAGTGGATTTGCAAAACCCCAATCCGTCAACTACCAATTTAAACCAACCACTGTTTCCA 367
Qy 181 AAATGGGGGTGATGCTTTGGGGTGGGGTGGGAAACCAACGGCTCTACCCCTACCGGTGT 240
Db 368 AAATGGGGGTGATGCTTTGGGGTGGGGTGGGAAACCAACGGCTCTACCCCTACCGGTGT 427
Qy 241 GTTATGCTTAACAGAGAGCAATTTATGCGCTACAAAGACACAGATTCAACAGCCAT 300
Db 428 GTTATGCGGAATCGAGAGGAATCTGCTGGGCCACAAAGGACACTGTGCACCAAGCCAT 487
Qy 301 TACTTTGGGCTCCCTACCCCAAGCCCTACGCTATTTGAGTTGAGTTCCTTCAGGAGAGGAA 360
Db 488 TATTTTGGTTCATTTACCCCAAGCATCAGCAATTAAGTTGGGCTTTACAAATGGAGAGAG 547
Qy 361 ATCTATGCCCATTCAGAGAGTCTGCTTCCAAATGTTTAAATCTGTACGACATTGTGTTGGG 420
Db 548 ATTTATGCTCATTCAGAGAGATCTCCCTATGGTGAACCAAGATGACATTGTGTTGGA 607
Qy 421 GGATGGATATCAGCAATGAACTGCGCTATGCTCAGTGGCCAGGGCAAAAGTGTGAC 480
Db 608 GGATGGGACATTAAGTACATGAACTTGCTGATGTCATGCGCAGGGCTTAGAGTGTATAT 667
Qy 481 ATCGATTTGCAAGAGCAGTTGAGGGCTTACATGGAATCATGTTTCCACTCCCGGGAATC 540
Db 668 ATTGATCTTCAAAAGCACTTAGGCTTTACATGGAATCAATGTTCCACTTCTGGAAATC 727
Qy 541 TACGACCCGGAATTCATTGCTGCGCAACCAAGAGAGCGTGCACCAACGTGATTAAGGGC 600
Db 728 TATGACCCGGAATTCATTGCTGCGCAATCAAGGGTGGCGTGCACCAACGTGATTAAGGA 787
Qy 601 ACAAGCAAGAGCAATTCAGCAATTCATCAAGACATCAAGGCTTTAAGGAAGCCACC 660
Db 788 ACCAAGAAAGAAACAATGCTCAAAATCATCAAGATATTAAGATTAAGAGAAAC 847
Qy 661 AAAGTGAACAAGGTGGTGTCTGTGAGCTGCGCAACAGAGAGTATAGCAATTTGGTT 720
Db 848 AAAGTGAATAGGGTGTGTGATCTGTGACCTGTACATCTGAGAGGTACGTAATGTGATT 907
Qy 721 GTAGGCTTAAATGACACCATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAG 780
Db 908 GTGGGGCTAAATGACACCATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAG 967
Qy 781 ATTTCTCTTCCACTTGTATGCCATTGCGCTGTGTGATGAGAAATGTTCTTTCAATAT 840
Db 968 ATTTCTCTTCCACTTGTATGCCATTGCGCTGTGTGTGTTGAAAATATCCCTTTCAATAT 1027
Qy 841 GGAACCCCTCAGAAACCTTTTGTACAGAGGCGGATGATCTTTGCCATGGCGAGAACCT 900
Db 1028 GGAACCCCTCAGAAACCTTTTGTACAGAGGCGGATGATCTTTGCCATGAGAGAACCT 1087
Qy 901 TTGATTTGGTGAATGACTTTCAAGAGTGTGTGACCAAAATGAATCTGTGTTGGTTGAT 960
Db 1088 TTGATTTGGTGAATGACTTTCAAGAGTGTGTGACCAAAATGAATCTGTGTTGGTTGAT 1147
Qy 961 TTTCTTTGGGGGGGTGTATCAAGGCAACATGTATATGTTAGTTACACCATTTGGGAAC 1020
Db 1148 TTTCTTTGGGGGGGTGTATCAAGGCAACATGTATATGTTAGTTAGTACACCATTTGGGAAC 1207
Qy 1021 AATGATGATGAATCTCGCGCTCCAAACCTTCGCTCCAGGAATCTCCAGAGC 1080
Db 1208 AATGATGATGAATCTCGCGCTCCAAACCTTCGCTCCAGGAATCTCCAGAGC 1267
Qy 1081 AACGTTTGTGAGATATGTTCAACAGCAATGCAATGCTCTATGAGCCTGTGTGAACATCC 1140
Db 1268 AATGTTTGTGAGATATGTTCAACAGCAATGCAATGCTCTATGAGCCTGTGTGAACATCC 1325
Qy 1141 GACCATGTTGTTGTTATTAATGATGTGCTTACGTAGGGGATAGCAAGAGACCATGGAT 1200

Db 1326 GACCACTGGTGG---ATCAAGTATGTGCCATATGTTGGC---ATCCAGAGAGGCCATGAGAT 1378
Qy 1201 GAGTACACTTCAGAGATATTCATGCGTGGAGAAAGAACACCATTTGTTTGCACAAACATGT 1260
Db 1379 GAGTACACATCAAGAAATTTTCATGGGTGGAGAGACCCATTTGCTGCGACAAACAGTGT 1438
Qy 1261 GAGGATTCCTTCTTACCTGTGCTCTATTTATCTTGACCTGGTCTTCTGTGTGAGTGGAG 1320
Db 1439 GAGGACTCTCTTTTGGCTGCTCCAAATCATTTCTAGATTTGGTTCTGCTGGCTGAATCAG 1498
Qy 1321 ACTAGATCCAGTTTAAAGCTGAAATGAGGAAATTCACACTCATTTCCACCCAGTTGCT 1380
Db 1499 ACCAGATTCAGCTCAAGCTGAGAGAGAGGCAATTCACCTTTCCACCTGTGGCT 1558
Qy 1381 ACCATTCTCAGCTATTTGACCAAGGCTCTCTGTTCACCGGGTACACCAAGTGTGAAT 1440
Db 1559 ACAATTTCTCAGCTACCTCAACAGGCTCTGCTGCTCCACAGGACACCAAGTGTGAAT 1618
Qy 1441 GCATTGTCAAGAGGCTGCAATGCTGGAATACATTAAGAGGCTTGTGTGATTTGGCC 1500
Db 1619 GGGCTGTCAAGAGGCTGCAATGCTTGAACATTTCTGAGAGGCTTGGCTTGGCT 1678
Qy 1501 CCAGAGATTAACATGATTTCTCGAGTACCAAGTGA 1533
Db 1679 CCTGAGAACACATGATTTTGAATACCAAGTGA 1711

RESULT 12
MCU32511 2053 bp mRNA linear PLN 11-JUL-1996
LOCUS Mesembryanthemum crystallinum myo-inositol-1-phosphate synthase
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Mesembryanthemum crystallinum (common iceplant)
Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Aizoaceae; Mesembryanthemum.

1 (bases 1 to 2053)
Ishitani, M., Majumder, A. L., Bornhouser, A., Michalowski, C. B.,
Jensen, R. G. and Bohner, H. J.

Coordinate transcriptional induction of myo-inositol metabolism
during environmental stress
Plant J. 9 (4), 537-548 (1996)

2 (bases 1 to 2053)
Ishitani, M., Majumder, A. L., Bornhouser, A., Michalowski, C. B.,
Jensen, R. G. and Bohner, H. J.

Direct Submission
Submitted (27-JUL-1995) Hans J. Bohner, Biochemistry, University
of Arizona, Biosciences West 516, Tucson, AZ 85721, USA

Location/Qualifiers

1. 2053
/organism="Mesembryanthemum crystallinum"
/mol_type="mRNA"
/db_xref="taxon:3544"
182..1720
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAB03687.1"
/translation="MFRISPKVESPNNVKTENBISVYNYDTTELHNRKQAGYOM
IVKPTVOYHFKTDTRVPLKGLVWLKGMGNNGSTLTGVIANRBSISATXDKIOAN
YRGSITQASIRVGSFNGSEIYAPKSLIPKPNPDVFGGDISDMNLADMTARV
FDIDQKQRPFMHNPVLPGLYDPDFLAANGSRANNVIGTKKEQVERVAKDIER
KEKNKRVVAVLTGNTERTSRVNVGLADTMENTLASKSESEISPSYALACIER
NIPIKNSPONTFVPLGLIDLAIKKNSLIGGDPFKSGTQKMSVLVDFVAGAIKPTSI
VSYNLGNNDGNLSAPOTPRSKETISKNVSDVAVASNGIIEYPEHDDHVVIVKYP
YVGSKRAMEYTSIEIFMGNTITVIMHNTCEBSLAAPIIIDVIVLALSTRIOAK

CDs

CDs

ORIGIN
BEDKFSHPHVATILSYLTKAPLVPGTPVNNALSKQRAMLENILBACVGLAENNMI
LETK"

Query Match 67.0%; Score 1027.4; DB 8; Length 2053;
Best Local Similarity 79.7%; Pred. No. 1.5e-269;
Matches 1227; Conservative 0; Mismatches 306; Indels 6; Gaps 1;

Qy 1 ATGTTATCGAGATTTTAAGTAGAGTCTTATGTGAAGTACCCAGACTGAGATT 60
Db 182 ATGTTATGAGAGCTTCAAGTAGAGACCACAAAGTGAAGTACAGAGATTA 241
Qy 61 CAGTCGTGTACAACTACGAACCCGGAACCTTTTCAAGAGAACGAA-----TGGC 114
Db 242 GAATCGGTGTACAACTACGAACCCGAGTTGGTTCAGAGAAATGCAAGATGCTGT 301
Qy 115 ACCATCACTGATTTGTCAAAACCCAAATCCGTCAACTCAATTTAAACCAACCCAT 174
Db 302 GGTATTCATGATCGTTTAAGCCCAAGACCTTCAGTATCACTTCAAGACTGACTGT 361
Qy 175 GTTCCAAATTTGGGGGTGATGCTTGTGGGTGGGAGAAACAGGCTTACCTCACC 234
Db 362 GTCCCAAACTGGGGGTGATGCTGTTGGATGGGGTGAATTAACGCTTCCACCTCA 421
Qy 235 GGTGGGTATTGCTTAACAGAGAGACATTTTATGGGCTTACAAAGAGATTCACAA 294
Db 422 GGTGGGTATGCTTCAACCGCAAGGAAATTTTGGGCAACCAAGAGATTCACGAA 481
Qy 295 GCCAATTACTTTGGCTCCCTCAACCCAGGCTCAGCTATTGAGTTGATCTTCCAGGA 354
Db 482 GCGAACTATTTGGTGTCTGTGACCCAGGCTTTTCAATTCGGGTGGGCTTTTAATGA 541
Qy 355 GAGAAATCTATGCTCCCATTTCAAGAGTCTGCTTCAATGTTTATCTGACGACAT 414
Db 542 GAGAGATTTATGCTCTTTCAAGAGCTTGCTCCATGTTGTAACCCAGATACGTA 601
Qy 415 TTTGGGGATGGGATATCAGCAACATGAACCTGGCTGATGCTATGGCCAGGGCAAG 474
Db 602 TTTGGGGGTGGGACATGAAGCAATGAACCTGGCTGATGACATGACCAAGGCTAG 661
Qy 475 TTTGACATGATTTTTCAGAGAGGTTGAGGCTTTCATGGAATTCATGTTTCACTGCC 534
Db 662 TTTGACATGATTTTTCAGAGAGGCTTTCATGGAATTCATGTTTCACTGCCCT 721
Qy 535 GGAATCTACGACCCGATTTTCAATTCGTCGCAACCAAGAGAGCGTCCCAACAGT 594
Db 722 GGTATTTACGACCTGATTTTCAATTCGTCGTCGCAACCAAGGCTCCGAGCAACG 781
Qy 595 AAGGCAACAAAGCAAGAGAGTTCAGCAATTCATCAAGACATCAAGGCTTTAAGGA 654
Db 782 AAGGCAACCAAGCAAGAGAGTTCAGCAATTCATCAAGACATTCAGGAGTTCAAG 841
Qy 655 GCCACCAAGTGAACAGAGTGTGCTTCTGTTGAGCTGCAACAGAGAGGATATGCAAT 714
Db 842 AAGAACAAAGTGAACAGAGTGTGCTTCTGTTGAGCTGCAACAGAGAGGATATG 901
Qy 715 TTTGTTAGAGCTTATGACACATGAGAGATCTTGTGCTGCTGTGAGACAGAAATG 774
Db 902 GTGTGTGTGGGCTCAAGACCATGAGAACCTGTTGGCATCTCTGAGAAAACAG 961
Qy 775 GTGAGATTTTCTCTTCACTGTTATGCTGCTGCTGCTGATGAGAAATGTTCTT 834
Db 962 TCGAGATTTTCTCTTCACTGTTATGCTGCTGCTGCTGATGAGAAATATTCCTT 1021
Qy 835 ATTAATGAGCCCTCAGAACCTTTTATCAAGGCTGATTTGATCTTGCCATGCGGAG 894
Db 1022 ATCAATGAGCCCTCAGAACCTTTTATCAAGGCTGATTTGATCTTGCCATGCGGAG 1081
Qy 895 AACACTTGAATTTGTGAGATGACTTCAAGAGTGTGACCAAAATGATCTGTGTTG 954
Db 1082 AACACTTGAATTTGTGAGATGACTTCAAGAGTGTGACCAAAATGATCTGTGTTG 1141
Qy 955 GTTGAATTTTCTTGGGGGCTGTGATCAAGCAACATCTATGTTAATCAACCATCTG 1014

Db 980 AAGGTAAGTCTCAGAAACCTTTGTCCTCCGTCTCAATGTAATGCGCAATGAAGAAT 1039
Qy 898 ACTTGAATTTGGTGAAGATGATCTTCAAGAGTGTGACCAAAATGAATCTGTGGT 957
Db 1040 TGTTTATTTGAGAGATGATGATCTTCAAGAGTGTGACCAATGAATGATGATGCT 1099
Qy 958 GATTTCTTGTGGGGGTGGTATCAAGCCCAATCTATAGTATGTAACAACCTGGGA 1017
Db 1100 GATTTCTTGTGGGGGTGGTATGAGCCCAATCTATGAGTATGACCACTTGGGA 1159
Qy 1018 AACCAATGATGATGATGATCTTCAAGAGTGTGACCAAACTTCCGCTCAAGAAATCTCCA 1077
Db 1160 AACCAATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTATGACCAAG 1219
Qy 1078 AGCAAGTGTGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1137
Db 1220 AGTAAATGATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1279
Qy 1138 CCCGACCATGTTGTGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1197
Db 1280 CCGATCAATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1339
Qy 1198 GATGATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1257
Db 1340 GATGATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1399
Qy 1258 TGTGAGATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1317
Db 1400 TGTGAGATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1459
Qy 1318 AGCACTGATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1377
Db 1460 AGCACTGATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1519
Qy 1378 GGTACATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1437
Db 1520 GGTACATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1579
Qy 1438 AATGATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1497
Db 1580 AATGATGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1636
Qy 1498 GCGGACGATGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1531
Db 1637 GCTCCAGAGATGATGATCTTCAAGAGTGTGACCACTTCAAGAGTGTGAA 1670

RESULT 14
AY028259 1845 bp mRNA linear PLN 26-MAR-2001
LOCUS Avicennia marina myo-inositol 1-phosphate synthase mRNA, complete
DEFINITION cds.
ACCESSION AY028259
VERSION AY028259.1 GI:13447456
KEYWORDS
SOURCE Avicennia marina
ORGANISM Avicennia marina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Lamiales; Acanthaceae; Acanthaceae incertae
sedis; Avicennia.
REFERENCE 1 (bases 1 to 1845)
AUTHORS Jithesh,M.N., Parami,M. and Parida,A.
TITLE Characterization of a cDNA for myo-inositol 1-phosphate synthase
from the mangrove species Avicennia marina
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1845)
JITHESH,M.N., PARAMI,M. and PARIDA,A.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-2001) Plant Molecular Biology Lab, M.S.
Swaminathan Research Foundation, III Cross Street, Chennai 600 113,
India

FEATURES
source
1..1845
/organism="Avicennia marina"
/mol_type="mRNA"
/db_xref="taxon:82927"
27..1556
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAK21969.1"
/db_xref="GI:13447457"
/translation="MPISPFKVDSPVKYTDSEIHSVNYETAELVHNRNGTYQIV
KPTVYKERTDTHVYKGLVLMVGMGNNGSTLTGTVIANKREGISWATQKQVAYF
GSLTQASAIRVGSFNGEIVAPFKSLIPVNVDDIIFGMDISNMMLADMGCGOLD
IDLOQRPRYMHMPPLPGIYYPDPIAAGSRANNVIGTKKQYQYIKDKERK
KTKVERVVGITANERYSNVVGLNDIASLWASLERNALISPSLVAIACLEDVP
FLNGSPONTFVGLIELAQRCLIGDDPKSGQTKMSVLVDPLVAGIKPTSPVSY
NHLGNDDMNLAPQTFRSKISKSVVDVMAVNSIILYEPESHPHVAVIKVPSVY
DKRAMDEVTSIEIFMGKSTIVHNTCESLAAPILDLVLAEISTRIOLKABEG
KPHSHFHVATILISYLTKALVLPPTPVVNAISKRAMLENILRACVGLAPENNMILEY
K"

ORIGIN
Query Match 66.3%; Score 1016.2; DB 8; Length 1845;
Best Local Similarity 79.4%; Pred. No. 1.7e-266;
Matches 1217; Conservative 0; Mismatches 313; Indels 3; Gaps 1;
Qy 1 ATGTTGATGAGAAATTTAAGGTAGAGAGTCTTAATGTGAAGTACACCGAGACTGAGATT 60
Db 27 ATGTTGATGAGAGGCTTCAAGGTGACAGCCCAAGTTAAGTACAGACAGTGAATT 86
Qy 61 CAGTCCGTGTACATCAGAAACCAACCGAATCTTTTCAAGAGAAAGAAATGACACTAT 120
Db 87 CAGTCCGTGTACATCAGAAACCAACCGAATCTTTTCAAGAGAAAGAAATGACACTAT 146
Qy 121 CAGTGAATGTCAAAACCCCAATCCGTCACATCAATTAATAACCAACCCCAATGTTCA 180
Db 147 CAGTGAATGTCAAGCCCAACCTGTTAATATGAGTTCAAGACCGAACCCATGTCGG 206
Qy 181 AAATGGGGGTATGCTTGGGGTGGGGTGAACCAACGGCTTACCTCAACCGGTGT 240
Db 207 AAATTAAGGTATATGCTTGGGGTGGGGGCAACAAATGTTCAACTGACAGGTGT 266
Qy 241 GTTATGCTCAACAGAGAGACATTTCAATGGGCTCAAAAGACAAATTTCAACAGCAT 300
Db 267 GTTATGCTCAACAGAGAGATTTCAATGGGCTCAAAAGACAAATTTCAACAGCAT 326
Qy 301 TACTTGGCTCCCTCAACCAACCTCAGCTATTCAGATGATGATCTTCAAGAGAGAA 360
Db 327 TACTTGGCTCCCTCAACCAACCTCAGCTATTCAGATGATGATCTTCAAGAGAGAA 386
Qy 361 ATCTATGCCCATTCAGAGATGCTTCCATGTTAATCTGACGACATGTGTTGG 420
Db 387 ATCTATGCTCCATTCAGAGATGCTTCCATGTTAATCTGACGACATGTGTTGG 446
Qy 421 GGATGGGATATAGCAACATGAACCTGCTGATGCTGACAGGCAAGAGGTGTTGAC 480
Db 447 GGATGGGATATAGCAACATGAATCTGCTGATGCTGACAGGCAAGAGGTGTTGAC 506
Qy 481 ATCGATTTGCAAGACAGTGAAGGCTTACATGGAATCAGATGCTTCCCGGGAATC 540
Db 507 ATCGATTTGCAAGACAGTGAAGGCTTACATGGAATCAGATGCTTCCCGGGAATC 566
Qy 541 TAGACCCGGATTTGATGCTGCTCCCAACCAAGAGAGCGTGCACCAACGCTGTTAAGG 600
Db 567 TACTACCTCGATCTTATGCTGCTCAATCAGGCTCAGCGGCCAACAAGTATCAAAG 626
Qy 601 ACAAGCAAGAGCAAGTTCAGCAATCTCAAGACATCAAGCGCTTTAAGAACCCACC 660
Db 627 ACAAGCAAGAGCAAGTTCAGCAATCTCAAGACATCAAGCGCTTTAAGAACCCACC 686
Qy 661 AAAGTGAACAAGGTGTTGCTGCTGACCTGACCAACAGAGAGGTATAGCAATTTGTT 720
Db 687 AAAGTGAACAAGGTGTTGCTGCTGACCTGACCAACAGAGAGGTATAGCAATTTGTT 743

QY	721	GTAGGCTTAATGACACCATGGAAATCTTGGCTGTGGACAGAAATGAGGCTGAG	780
DB	744	GTGGGCTTAATGATACACGAGAAACCTTAATGCTCTCTGGAAGAAATGAGCTGAG	803
QY	781	ATTCTCTCTTCAACCTTGTATGCAATGCTGTGTGATGAAAATGTTCTTTCAATAT	840
DB	804	ATATCTCTTCAACCTTATATGCAATGCTGTGTGATGAAAATGTTCTTTCAATAT	863
QY	841	GGAAAGCCCTCAGAAACCTTTGTACAGAGGCTGATGATCTTGCCATGCCGAGAACT	900
DB	864	GGCAGCCCAAAACCTTTCCTTCCAGGCTGATGATGATGATGATGATGATGATGAT	923
QY	901	TGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	960
DB	924	TGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	983
QY	961	TTTCTTGTGGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	1020
DB	984	TTCTCTTGTGGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	1043
QY	1021	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1080
DB	1044	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1103
QY	1081	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1140
DB	1104	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1163
QY	1141	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1200
DB	1164	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1223
QY	1201	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1260
DB	1224	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1283
QY	1261	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1320
DB	1284	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1343
QY	1321	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1380
DB	1344	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1403
QY	1381	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1440
DB	1404	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1463
QY	1441	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1500
DB	1464	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1523
QY	1501	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1533
DB	1524	TAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1556
RESULT 15			
LOCUS	AX054630	1781 bp	DNA
DEFINITION	Sequence 1 from Patent WO0073473.	linear	PAT 13-JAN-2001
ACCESSION	AX054630		
VERSION	AX054630.1	GI:12228190	
KEYWORDS			
SOURCE	Brassica napus (rape)		
ORGANISM	Brassica napus		
REFERENCE	1	Georges, F., Hussain, A.A. and Keller, W.A.	
AUTHORS			
TITLE	Method for reducing phytate in canola meal using genetic		

JOURNAL	manipulation involving myo-inositol 1-phosphate synthase gene									
PATENT	Patent: WO 0073473-A 1 07-DEC-2000;									
NATIONAL	RESEARCH COUNCIL CANADA (CA)									
FEATURES	Location/Qualifiers									
SOURCE	1. 1781									
	/organism="Brassica napus"									
	/mol_type="unassigned DNA"									
	/db_xref="taxon:3708"									
ORIGIN										
Query Match	65.2%;	Score 1000.2;	DB 6;	Length 1781;						
Best Local Similarity	78.3%;	Pred. No. 4.2e-262;								
Matches 1200;	Conservative 0;	Mismatches 333;	Indels 0;	Gaps 0;						
QY	1	ATGTCATGAGAAATTTTAAAGTATGAGAGTCTTAATGAGATACCGAGACTGAGATT	60							
DB	57	ATGTCATGAGAGACTTTCAAAGTCGAGAGCCCGAACGGAATGACACCGAATGAGATT	116							
QY	61	CAGTCGCTGATCACTGTCGAAACACCGAATCTTGTCACGGAACAGGAATGGACCAT	120							
DB	117	CATTGGGTGATGATTCAGAGACACGAGGTGCTTACAGGAACGTCAACGGTCTTAC	176							
QY	121	CAGTGAATGTCGAACCCCAATCCGTCACATCACAAATTTAAACCAACCCATGTTCCA	180							
DB	177	CAGTGAATCGGAAGCCCAAGTGTGTCAAAATACATTTCAAAACCGACACTGTGTCCC	236							
QY	181	AAATTTGGGGGTGATGCTTGTGGTGGTGGTGGAAACAACGGCTTACCTTCAACGGTGT	240							
DB	237	AAATTTAGGGGTATGCTTGTGTGGTGGGAGGAACAATGATCAACCTTCAACCGCTGT	296							
QY	241	GTATTTGCTAACAGAGGAGCATTTCAATGGGCTCAAAAGCAAGATTCAACAAGCCAT	300							
DB	297	GTAATTTGCATTAAGAGGAATATCTCGGGCGACCAAAGACAAAGTGCACAAACGAC	356							
QY	301	TACTTTGGCTCTCTCACCCCAAGCTCAGCTATTGAGTTGATCTTTCAGGGAGAGAA	360							
DB	357	TACTTCGGGTGTTAACACAGCATCGCTATTTGCTGTGCGATCTTTAAACGTGTAAG	416							
QY	361	ATCTATGCCCATTCAGAGTCTGCTTCCATATGTTAACTCTGACGACATGTGTTGGG	420							
DB	417	ATGATGCCCCCTTCAAGAGTCTGTTCCATATGGGAATCCGGAATGATGTGTGTTGGA	476							
QY	421	GGATGGGATACACCAACATGAACCTGGCTGATGCAATGGCCAGGGCAAAAGTGTTAC	480							
DB	477	GGATGGGACATTAACGATGAATTAATTTAGCAACGAGTGGGTAGAGCCAAAGTCTTGA	536							
QY	481	ATCGATTTGCAAGAGAGTGTAGGCTCTTACATGGAATCATGTGTTCCACTCCCGGAAT	540							
DB	537	ATTGATCTGCAAGAAACAGCTACGGCTTACATGAGAAACATTTGTCCCACTCCCTGGAT	596							
QY	541	TACGACCCGGATTCATTGCTGCCAACCAAGAGAGCTGCCAACACGTGATTAAAGGC	600							
DB	597	TACGACCCGTGATTCATGCTGCCAATTCAGAGCTCACCTGCCAACACGTATCAAAAG	656							
QY	601	ACAAAGCAAGGCAAGTTCACGAATATCTAAAGACATCAAGGGGTTTAAAGAACCCAC	660							
DB	657	ACCAAGAAAGCAAGCTGCACCAATCATCAAGACATGAGAGGAGTTTAAAGAGAAAC	716							
QY	661	AAAGTGAAGAGTGTGTCTCTGTGACCTGCCAAACAGAGAGGATATAGCAATTTGGTT	720							
DB	717	AAAGTGAATAGGTTGTGTGTCTGTGACGGCTTAACAGAGCGTTTACGCAATGTGAT	776							
QY	721	GTAGGCTTAATGACACCATGGAGAAATCTTGGCTGTGTGACAGAAATGAGGCTGAG	780							
DB	777	GTGGGGCTTAACGACATATGAGAAATCTTATGAACCTGTGTGATAGGGAATGAGCTGAG	836							
QY	781	ATTTCCTCTTCAACCTGTATGCAATGCTGTGTGATGAAAATGTTCTTTCAATAT	840							
DB	837	ATTTCCTCTTCAACGCTTATGCAATGCAATGTGTTCTTGAAGATATCTTTCAATAT	896							
QY	841	GGAAAGCCCTCAGAAACCTTTGTACAGAGGCTGATGATCTTTGCCATGGCGAGAACCT	900							
DB	897	GGAAAGCCCTCAGAAACCTTTGTGTCCGGGCTTATATGATTTGGCTATACAGACATGTT	956							

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 16:46:53 ; Search time 81.5 Seconds
(without alignments)
2254.032 Million cell updates/sec

Title: US-10-718-952-2

Perfect score: 2632
Sequence: 1 MFIEHFKECPNVKXTEREI.....NIMRACVGLAPENNMLEYK 510

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1599520 seqs, 360203123 residues

Total number of hits satisfying chosen parameters: 1599520

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pdb.*
2: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pdb.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pdb.*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pdb.*
5: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pdb.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pdb.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pdb.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pdb.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pdb.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pdb.*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pdb.*
12: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pdb.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pdb.*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pdb.*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pdb.*
16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pdb.*
17: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pdb.*
18: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pdb.*
19: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pdb.*
20: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pdb.*
21: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2632	100.0	510	US-10-025-003-2	Sequence 2, Appl1
2	2632	100.0	510	US-10-025-003-10	Sequence 10, Appl1
3	2632	100.0	510	US-10-718-952-2	Sequence 2, Appl1
4	2632	100.0	510	US-10-718-952-10	Sequence 10, Appl1
5	2627	99.8	510	US-10-025-003-6	Sequence 6, Appl1
6	2627	99.8	510	US-10-718-952-6	Sequence 6, Appl1
7	2621.5	99.6	511	US-10-424-599-213009	Sequence 213009,
8	2600	98.8	510	US-10-025-003-14	Sequence 14, Appl1
9	2600	98.8	510	US-10-025-003-16	Sequence 16, Appl1
10	2600	98.8	510	US-10-718-952-14	Sequence 14, Appl1
11	2600	98.8	510	US-10-718-952-16	Sequence 16, Appl1
12	2593	98.5	510	US-10-025-003-12	Sequence 12, Appl1

13	2593	98.5	510	US-10-718-952-12	Sequence 12, Appl1
14	2524	95.9	510	US-10-424-599-154864	Sequence 154864,
15	2432	92.4	505	US-10-442-017-15	Sequence 15, Appl1
16	2356	89.5	510	US-10-425-115-231852	Sequence 231852,
17	2356	89.5	516	US-10-425-114-62568	Sequence 62568, A
18	2355	89.5	510	US-10-767-701-65278	Sequence 46278, A
19	2354	89.4	510	US-10-739-930-7635	Sequence 7635, Ap
20	2351	89.3	510	US-10-425-115-306814	Sequence 306814,
21	2347	89.2	510	US-10-437-963-173252	Sequence 173252,
22	2345	89.1	542	US-10-425-114-58674	Sequence 58674, A
23	2336	88.8	510	US-09-921-232-11	Sequence 11, Appl1
24	2336	88.8	510	US-09-921-330-11	Sequence 11, Appl1
25	2336	88.8	510	US-09-921-329-11	Sequence 11, Appl1
26	2335	88.7	510	US-09-727-628-2	Sequence 2, Appl1
27	2335	88.7	510	US-10-425-115-231850	Sequence 231850,
28	2335	88.7	510	US-10-425-115-231857	Sequence 231857,
29	2335	88.7	512	US-10-425-114-59621	Sequence 39621, A
30	2335	88.7	512	US-10-425-114-53231	Sequence 53231, A
31	2335	88.7	512	US-10-425-114-66216	Sequence 66216, A
32	2334	88.7	510	US-10-425-115-231853	Sequence 231853,
33	2334	88.7	512	US-10-425-114-46916	Sequence 46916, A
34	2205.5	83.8	509	US-10-437-963-187768	Sequence 187768,
35	2193	83.3	510	US-10-425-115-268216	Sequence 268216,
36	2193	83.3	521	US-10-425-114-68859	Sequence 68859, A
37	2193	83.3	522	US-10-425-114-62888	Sequence 62888, A
38	2193	83.3	529	US-10-425-114-58895	Sequence 58895, A
39	2129	80.9	458	US-10-425-115-231843	Sequence 231843,
40	2104	79.9	450	US-10-425-115-231845	Sequence 231845,
41	2016	76.6	431	US-10-424-599-154863	Sequence 154863,
42	1607	61.1	558	US-09-876-889-348	Sequence 348, App
43	1596.5	60.7	534	US-10-128-714-3515	Sequence 3515, Ap
44	1592.5	60.5	534	US-10-128-714-8515	Sequence 8515, Ap
45	1558	59.2	332	US-10-425-114-42702	Sequence 42702, A

ALIGNMENTS

RESULT 1
US-10-025-003-2
; Sequence 2, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025.003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-025-003-2
Query Match 100.0%; Score 2632; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 8.6e-220; Mismatches 0; Indels 0; Gaps 0;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFIEHFKECPNVKXTEREIQSVNYETTELVHNRNQTGYIVKPKSVKYEKFNTHVP 60
DB 1 MFIEHFKECPNVKXTEREIQSVNYETTELVHNRNQTGYIVKPKSVKYEKFNTHVP 60
QY 61 KLGVVLVWGNGNSTLTGVIARREGISMATKDKIQQANTFGSLTQSAIRVGSFQEE 120

```
Db      61 KLGWMLVGMGNGNSTLTGVIANREGISWATKDKIQANYPGSLTQASAIRVGSFOGEE 120
Qy      121 IYAPFKSLPWNPPDDIYFGWMDISNMNLADAMARAKVFDIDLOQLRPYMESMLPLPGI 180
Db      121 IYAPFKSLPWNPPDDIYFGWMDISNMNLADAMARAKVFDIDLOQLRPYMESMLPLPGI 180
Qy      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIAKFKATKVDKVVVMTANTERYSNLV 240
Db      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIAKFKATKVDKVVVMTANTERYSNLV 240
Qy      241 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIAKFKATKVDKVVVMTANTERYSNLV 240
Db      241 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIAKFKATKVDKVVVMTANTERYSNLV 240
Qy      301 LIIGDDFKSGQTKMSVLDPLVAGIKPTISVSYNHLGNNDGMNLSAPQFRSKSEIKS 360
Db      301 LIIGDDFKSGQTKMSVLDPLVAGIKPTISVSYNHLGNNDGMNLSAPQFRSKSEIKS 360
Qy      361 NVDDMVNSNALIYEPGEHPDHVVVIXVYPVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Db      361 NVDDMVNSNALIYEPGEHPDHVVVIXVYPVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy      421 EDSLAAPIIIDLVLLAELSTRIEFKANEKGKFSFHPVATILSYLTKAPLVPPGTPVNV 480
Db      421 EDSLAAPIIIDLVLLAELSTRIEFKANEKGKFSFHPVATILSYLTKAPLVPPGTPVNV 480
Qy      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 2

```
US-10-025-003-10
; Sequence 10, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025, 003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-10
```

```
Query Match      100.0%; Score 2632; DB 14; Length 510;
Best Local Similarity 100.0%; Pred. No. 8.6e-220;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 MFIENFKYECPNVKYKTEIETIOSVYNYETTEL VHENRNGTYOMIVKPKSVKTEFKTNIHVP 60
Db      1 MFIENFKYECPNVKYKTEIETIOSVYNYETTEL VHENRNGTYOMIVKPKSVKTEFKTNIHVP 60
Qy      61 KLGWMLVGMGNGNSTLTGVIANREGISWATKDKIQANYPGSLTQASAIRVGSFOGEE 120
Db      61 KLGWMLVGMGNGNSTLTGVIANREGISWATKDKIQANYPGSLTQASAIRVGSFOGEE 120
Qy      121 IYAPFKSLPWNPPDDIYFGWMDISNMNLADAMARAKVFDIDLOQLRPYMESMLPLPGI 180
Db      121 IYAPFKSLPWNPPDDIYFGWMDISNMNLADAMARAKVFDIDLOQLRPYMESMLPLPGI 180
Qy      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIAKFKATKVDKVVVMTANTERYSNLV 240
Db      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIAKFKATKVDKVVVMTANTERYSNLV 240
Qy      241 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIAKFKATKVDKVVVMTANTERYSNLV 240
Db      241 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIAKFKATKVDKVVVMTANTERYSNLV 240
Qy      301 LIIGDDFKSGQTKMSVLDPLVAGIKPTISVSYNHLGNNDGMNLSAPQFRSKSEIKS 360
Db      301 LIIGDDFKSGQTKMSVLDPLVAGIKPTISVSYNHLGNNDGMNLSAPQFRSKSEIKS 360
Qy      361 NVDDMVNSNALIYEPGEHPDHVVVIXVYPVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Db      361 NVDDMVNSNALIYEPGEHPDHVVVIXVYPVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy      421 EDSLAAPIIIDLVLLAELSTRIEFKANEKGKFSFHPVATILSYLTKAPLVPPGTPVNV 480
Db      421 EDSLAAPIIIDLVLLAELSTRIEFKANEKGKFSFHPVATILSYLTKAPLVPPGTPVNV 480
Qy      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db      481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 3

```
US-10-718-952-2
; Sequence 2, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718, 952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-2
```

```
Query Match      100.0%; Score 2632; DB 16; Length 510;
Best Local Similarity 100.0%; Pred. No. 8.6e-220;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 MFIENFKYECPNVKYKTEIETIOSVYNYETTEL VHENRNGTYOMIVKPKSVKTEFKTNIHVP 60
Db      1 MFIENFKYECPNVKYKTEIETIOSVYNYETTEL VHENRNGTYOMIVKPKSVKTEFKTNIHVP 60
Qy      61 KLGWMLVGMGNGNSTLTGVIANREGISWATKDKIQANYPGSLTQASAIRVGSFOGEE 120
Db      61 KLGWMLVGMGNGNSTLTGVIANREGISWATKDKIQANYPGSLTQASAIRVGSFOGEE 120
Qy      121 IYAPFKSLPWNPPDDIYFGWMDISNMNLADAMARAKVFDIDLOQLRPYMESMLPLPGI 180
Db      121 IYAPFKSLPWNPPDDIYFGWMDISNMNLADAMARAKVFDIDLOQLRPYMESMLPLPGI 180
Qy      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIAKFKATKVDKVVVMTANTERYSNLV 240
Db      181 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIAKFKATKVDKVVVMTANTERYSNLV 240
Qy      241 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIAKFKATKVDKVVVMTANTERYSNLV 240
Db      241 YDPDFIAANOEBRANNVIKGTQOEVOQI IKDIAKFKATKVDKVVVMTANTERYSNLV 240
```

```

Db 241 VGLNDTMENLAAVDRNEAIEISPTLYALACVMEVPIFGSPQNTFVGLIDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYNHLGNDGMNLSAQPTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYNHLGNDGMNLSAQPTFRSKEISKS 360
Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Qy 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPPVN 480
Db 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPPVN 480
Qy 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510

```

RESULT 4

```

US-10-718-952-10
; Sequence 10, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strelt, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-10

```

Query Match 100.0%; Score 2632; DB 16; Length 510;
 Best Local Similarity 100.0%; Pred. No. 8.6e-220;
 Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MFIENFKVCEPNVKTETETIQSVNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Db 1 MFIENFKVCEPNVKTETETIQSVNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Qy 61 KLGWLVGVGNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGGE 120
Db 61 KLGWLVGVGNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGGE 120
Qy 121 IYAPFKSLPVMNPDPDIYFGWMDISNMNLADAMAAKVPDIDLOQLRPFYMSMLPLPGI 180
Db 121 IYAPFKSLPVMNPDPDIYFGWMDISNMNLADAMAAKVPDIDLOQLRPFYMSMLPLPGI 180
Qy 121 IYAPFKSLPVMNPDPDIYFGWMDISNMNLADAMAAKVPDIDLOQLRPFYMSMLPLPGI 180
Db 121 IYAPFKSLPVMNPDPDIYFGWMDISNMNLADAMAAKVPDIDLOQLRPFYMSMLPLPGI 180
Qy 181 YDPDFIAANQEBRANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERTSNLY 240
Db 181 YDPDFIAANQEBRANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERTSNLY 240
Qy 181 YDPDFIAANQEBRANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERTSNLY 240
Db 181 YDPDFIAANQEBRANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERTSNLY 240
Qy 241 VGLNDTMENLAAVDRNEAIEISPTLYALACVMEVPIFGSPQNTFVGLIDLAIARNT 300
Db 241 VGLNDTMENLAAVDRNEAIEISPTLYALACVMEVPIFGSPQNTFVGLIDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYNHLGNDGMNLSAQPTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYNHLGNDGMNLSAQPTFRSKEISKS 360

```

```

Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Qy 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPPVN 480
Db 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPPVN 480
Qy 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510

```

RESULT 5

```

US-10-025-003-6
; Sequence 6, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Strelt, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-025-003-6

```

Query Match 99.8%; Score 2627; DB 14; Length 510;
 Best Local Similarity 99.8%; Pred. No. 2.3e-219;
 Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 MFIENFKVCEPNVKTETETIQSVNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Db 1 MFIENFKVCEPNVKTETETIQSVNYETTELVEHNRNGTYQWIVKPKSVKYEFTKNIHVP 60
Qy 61 KLGWLVGVGNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGGE 120
Db 61 KLGWLVGVGNGNSTLTGCVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFGGE 120
Qy 121 IYAPFKSLPVMNPDPDIYFGWMDISNMNLADAMAAKVPDIDLOQLRPFYMSMLPLPGI 180
Db 121 IYAPFKSLPVMNPDPDIYFGWMDISNMNLADAMAAKVPDIDLOQLRPFYMSMLPLPGI 180
Qy 121 IYAPFKSLPVMNPDPDIYFGWMDISNMNLADAMAAKVPDIDLOQLRPFYMSMLPLPGI 180
Db 121 IYAPFKSLPVMNPDPDIYFGWMDISNMNLADAMAAKVPDIDLOQLRPFYMSMLPLPGI 180
Qy 181 YDPDFIAANQEBRANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERTSNLY 240
Db 181 YDPDFIAANQEBRANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERTSNLY 240
Qy 181 YDPDFIAANQEBRANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERTSNLY 240
Db 181 YDPDFIAANQEBRANNVIKGTQEOVOQIIKDIKAFKATKVDKVVVLTANTERTSNLY 240
Qy 241 VGLNDTMENLAAVDRNEAIEISPTLYALACVMEVPIFGSPQNTFVGLIDLAIARNT 300
Db 241 VGLNDTMENLAAVDRNEAIEISPTLYALACVMEVPIFGSPQNTFVGLIDLAIARNT 300
Qy 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYNHLGNDGMNLSAQPTFRSKEISKS 360
Db 301 LIGGDDFKSGQTKMSVLDVFLVAGIKPTSIYSYNHLGNDGMNLSAQPTFRSKEISKS 360
Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVVKYVPYVGDSCRAMDEYTSSEIFMGKSTIYLAHNTC 420
Qy 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPPVN 480

```

```
Db 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKALVPPTGVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 6

US-10-718-952-6

```
/ Sequence 6, Application US/10718952
/ Publication No. US20040128713A1
/ GENERAL INFORMATION:
/ APPLICANT: Hitz, William
/ APPLICANT: Sebastian, Scott
/ APPLICANT: Grace, John
/ APPLICANT: Streitz, Leon
/ TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
/ TITLE OF INVENTION: SACCCHARIDES AND PHYTIC ACID
/ FILE REFERENCE: BB-1077-C
/ CURRENT APPLICATION NUMBER: US/10/718, 952
/ PRIOR FILING DATE: 2003-11-21
/ PRIOR APPLICATION NUMBER: 08/835, 751
/ PRIOR FILING DATE: APRIL 8, 1997
/ PRIOR APPLICATION NUMBER: PCT/US98/06822
/ PRIOR FILING DATE: APRIL 7, 1998
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 6
/ LENGTH: 510
/ TYPE: PRT
/ ORGANISM: Glycine max
US-10-718-952-6
```

```
Query Match 99.8%; Score 2627; DB 16; Length 510;
Best Local Similarity 99.8%; Pred. No. 2,3e-219;
Matches 509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 MFIENFKVCEPNVKTETETESVYNYVETTELVHNRNGTYOMIVKPKSVKTEFKTNHVP 60
Db 1 MFIENFKVCEPNVKTETETESVYNYVETTELVHNRNGTYOMIVKPKSVKTEFKTNHVP 60
Qy 61 KLGVMLVGMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGE 120
Db 61 KLGVMLVGMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGE 120
Qy 121 IYAPFKSLIPVNPDDIVFGGWDISNMNLADAMARAKVPDIDLOKQLRPYMESMLPLPG 180
Db 121 IYAPFKSLIPVNPDDIVFGGWDISNMNLADAMARAKVPDIDLOKQLRPYMESMLPLPG 180
Qy 181 YPDPFIANOEBERANNVIGTKQEOVOQIIKDIAFKKATKVDKVVVLTANTERYSNLV 240
Db 181 YPDPFIANOEBERANNVIGTKQEOVOQIIKDIAFKKATKVDKVVVLTANTERYSNLV 240
Qy 241 VGLNDTMENLLAAVDRNEAEISPSSTLYAIACVMENVPFINGSPONTFVPGILDLAIARN 300
Db 241 VGLNDTMENLLAAVDRNEAEISPSSTLYAIACVMENVPFINGSPONTFVPGILDLAIARN 300
Qy 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNLSAPQTFRSKEISK 360
Db 301 LIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNLSAPQTFRSKEISK 360
Qy 361 NVVDVMNNSNALIYEPGEHPDHVVVIVKYPVYGDSCRAMDEYTSSEIFMGKSTIVLHNT 420
Db 361 NVVDVMNNSNALIYEPGEHPDHVVVIVKYPVYGDSCRAMDEYTSSEIFMGKSTIVLHNT 420
Qy 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKALVPPTGVN 480
Db 421 EDSLAAPIILDLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKALVPPTGVN 480
Qy 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
```

RESULT 7

US-10-424-599-213009

```
/ Sequence 213009, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424, 599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 213009
/ LENGTH: 511
/ TYPE: PRT
/ ORGANISM: Glycine max
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(511)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1.pep
US-10-424-599-213009
```

```
Query Match 99.6%; Score 2621.5; DB 15; Length 511;
Best Local Similarity 99.8%; Pred. No. 7e-219;
Matches 510; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
Qy 1 MFIENFKVCEPNVKTETETESVYNYVETTELVHNRNGTYOMIVKPKSVKTEFKTNHVP 60
Db 1 MFIENFKVCEPNVKTETETESVYNYVETTELVHNRNGTYOMIVKPKSVKTEFKTNHVP 60
Qy 61 KLGVMLVGMG-GNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGE 119
Db 61 KLGVMLVGMGNGNSTLTGGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFQGE 120
Qy 120 EYAPFKSLIPVNPDDIVFGGWDISNMNLADAMARAKVPDIDLOKQLRPYMESMLPLPG 179
Db 121 EYAPFKSLIPVNPDDIVFGGWDISNMNLADAMARAKVPDIDLOKQLRPYMESMLPLPG 180
Qy 180 IYDPDFIANOEBERANNVIGTKQEOVOQIIKDIAFKKATKVDKVVVLTANTERYSNLV 239
Db 181 IYDPDFIANOEBERANNVIGTKQEOVOQIIKDIAFKKATKVDKVVVLTANTERYSNLV 240
Qy 240 VGLNDTMENLLAAVDRNEAEISPSSTLYAIACVMENVPFINGSPONTFVPGILDLAIARN 299
Db 241 VGLNDTMENLLAAVDRNEAEISPSSTLYAIACVMENVPFINGSPONTFVPGILDLAIARN 300
Qy 300 TLIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNLSAPQTFRSKEISK 359
Db 301 TLIGGDDFKSGQTKMKSVLVDFLVGAGIKPTSIYSYNHLGNNDGNLSAPQTFRSKEISK 360
Qy 360 SNVVDVMNNSNALIYEPGEHPDHVVVIVKYPVYGDSCRAMDEYTSSEIFMGKSTIVLHNT 419
Db 361 SNVVDVMNNSNALIYEPGEHPDHVVVIVKYPVYGDSCRAMDEYTSSEIFMGKSTIVLHNT 420
Qy 420 CEDSLAAPIILDLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKALVPPTGVN 479
Db 421 CEDSLAAPIILDLVLAELSTRIEFKAENEGKFSHPVATILSYLTAKALVPPTGVN 480
Qy 480 NALSKORAMLENIMRACVGLAPENNMILEYK 510
Db 481 NALSKORAMLENIMRACVGLAPENNMILEYK 511
```

RESULT 8

US-10-025-003-14

/ Sequence 14, Application US/10025003


```

; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-025-003-14

Query Match          98.8%; Score 2600; DB 14; Length 510;
Best Local Similarity 98.6%; Pred. No. 5, 2e-217;
Matches 503; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MEIENKVCSPVVKYETEIQSVNYVETTELVEHNRNGTYQWIVKPKSVKYEKFTNIHVP 60
Db 1 MEIENKVCSPVVKYETEIQSVNYVETTELVEHNRNGTYQWIVKPKSVKYEKFTNIHVP 60
Qy 61 KIGWMLVGVGNNGSTLTGVIANREGISMATKDKIQOANYGSLTQASAIRVGSFGGEE 120
Db 61 KIGWMLVGVGNNGSTLTGVIANREGISMATKDKIQOANYGSLTQASAIRVGSFGGEE 120
Qy 121 IYAPFSLPMTVPNDIVFGMDISNMNLADAMARAKVFDIDLOKOLRPMESMLPLPGI 180
Db 121 IYAPFSLPMTVPNDIVFGMDISNMNLADAMARAKVFDIDLOKOLRPMESMLPLPGI 180
Qy 181 YPDPFIAANQEBRANNVIKGTQEQVOQIIKDIKAFKATKVDKVVLTANTATERSNLV 240
Db 181 YPDPFIAANQEBRANNVIKGTQEQVOQIIKDIKAFKATKVDKVVLTANTATERSNLV 240
Qy 241 VGLNDTMENTLLAADVBNENAEISPTLYAIAACWENVPFNGSPQNTFVGLIDLAIART 300
Db 241 VGLNDTMENTLLAADVBNENAEISPTLYAIAACWENVPFNGSPQNTFVGLIDLAIART 300
Qy 301 LIGGDDPKSGQTKMSVLVDPLVAGIKPTSIYSYHNLGNNDGMNLSAPQTRSKSISKS 360
Db 301 LIGGDDPKSGQTKMSVLVDPLVAGIKPTSIYSYHNLGNNDGMNLSAPQTRSKSISKS 360
Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVIKVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLLAAPITLIDLVLAEISTRIEKAENEGKFSHPVATILSYLTKAPLVPPGTPVYN 480
Db 421 EDSLLAAPITLIDLVLAEISTRIEKAENEGKFSHPVATILSYLTKAPLVPPGTPVYN 480
Qy 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510

RESULT 9
US-10-025-003-16
; Sequence 16, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon

```

```

; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-025-003-16

Query Match          98.8%; Score 2600; DB 14; Length 510;
Best Local Similarity 98.6%; Pred. No. 5, 2e-217;
Matches 503; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MEIENKVCSPVVKYETEIQSVNYVETTELVEHNRNGTYQWIVKPKSVKYEKFTNIHVP 60
Db 1 MEIENKVCSPVVKYETEIQSVNYVETTELVEHNRNGTYQWIVKPKSVKYEKFTNIHVP 60
Qy 61 KIGWMLVGVGNNGSTLTGVIANREGISMATKDKIQOANYGSLTQASAIRVGSFGGEE 120
Db 61 KIGWMLVGVGNNGSTLTGVIANREGISMATKDKIQOANYGSLTQASAIRVGSFGGEE 120
Qy 121 IYAPFSLPMTVPNDIVFGMDISNMNLADAMARAKVFDIDLOKOLRPMESMLPLPGI 180
Db 121 IYAPFSLPMTVPNDIVFGMDISNMNLADAMARAKVFDIDLOKOLRPMESMLPLPGI 180
Qy 181 YPDPFIAANQEBRANNVIKGTQEQVOQIIKDIKAFKATKVDKVVLTANTATERSNLV 240
Db 181 YPDPFIAANQEBRANNVIKGTQEQVOQIIKDIKAFKATKVDKVVLTANTATERSNLV 240
Qy 241 VGLNDTMENTLLAADVBNENAEISPTLYAIAACWENVPFNGSPQNTFVGLIDLAIART 300
Db 241 VGLNDTMENTLLAADVBNENAEISPTLYAIAACWENVPFNGSPQNTFVGLIDLAIART 300
Qy 301 LIGGDDPKSGQTKMSVLVDPLVAGIKPTSIYSYHNLGNNDGMNLSAPQTRSKSISKS 360
Db 301 LIGGDDPKSGQTKMSVLVDPLVAGIKPTSIYSYHNLGNNDGMNLSAPQTRSKSISKS 360
Qy 361 NVVDDMVNSNALIYEPGEHPDHVVVIKVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVVDDMVNSNALIYEPGEHPDHVVVIKVPYVGDSCRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLLAAPITLIDLVLAEISTRIEKAENEGKFSHPVATILSYLTKAPLVPPGTPVYN 480
Db 421 EDSLLAAPITLIDLVLAEISTRIEKAENEGKFSHPVATILSYLTKAPLVPPGTPVYN 480
Qy 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510
Db 481 ALSKORAMLENTIRACVGLAPENNMILEYK 510

RESULT 10
US-10-718-952-14
; Sequence 16, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIOR FILING DATE: 2003-11-21

```

;; PRIOR FILING DATE: APRIL 8, 1997
;; PRIOR APPLICATION NUMBER: PCT/US98/06822
;; PRIOR FILING DATE: APRIL 7, 1998
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 14
;; LENGTH: 510
;; TYPE: PRT
;; ORGANISM: Glycine max
US-10-718-952-14

Query Match 98.8%; Score 2600; DB 16; Length 510;
Best Local Similarity 98.6%; Pred. No. 5.2e-217;
Matches 503; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY	1	MEIENFKVCEPNVKTETETI	QSVNYVNETTEL	VHENRNGTYQMI	VKPKSVKEFKNTINHP	60
DB	1	MEIENFKVCEPNVKTETETI	QSVNYVNETTEL	VHENRNGTYQMI	VKPKSVNYQFKNTINHP	60
QY	61	KLGVMLVGMGNGNSTLTG	VIANREGISWATKDKI	QOANYFGSLTQASAIRVGS	FOGEE	120
DB	61	KLGVMLVGMGNGNSTLTG	VIANREGISWATKDKI	QOANYFGSLTQASAIRVGS	FOGEE	120
QY	121	IYAPFKSLIPMNPPD	IVFGGWDISNMNLADMA	RAKVFDDLOKOLRPYMESMLPLPGI	180	
DB	121	IYAPFKSLIPMNPPD	IVFGGWDISNMNLADMA	RAKVFDDLOKOLRPYMESMLPLPGI	180	
QY	181	YDPDFIAANOEERANNVI	KGTKOEVOQI	IKDIAKFKATKVDKVVVLTANTERYSNLV	240	
DB	181	YDPDFIAANOEERANNVI	KGTKOEVOQI	IKDIAKFKATKVDKVVVLTANTERYSNLV	240	
QY	241	VGNDTMENLLAADVDRNEA	ISPSSTLYAIACWENVP	FINSGPONTFVPGILDLAIANT	300	
DB	241	VGNDTMENLLAADVDRNEA	ISPSSTLYAIACWENVP	FINSGPONTFVPGILDLAIANT	300	
QY	301	LIGGDDFKSGQTKMSV	VDPLVAGIKPTSTIVSYNH	GNNDGMLSAPOFRSKEISKS	360	
DB	301	LIGGDDFKSGQTKMSV	VDPLVAGIKPTSTIVSYNH	GNNDGMLSAPOFRSKEISKS	360	
QY	361	NVDDMVNSNALIYEPGE	HPHVVIKYVPYVGS	KRAMDEYTSSEIFMGKSTIYLAHNTC	420	
DB	361	NVDDMVNSNALIYEPGE	HPHVVIKYVPYVGS	KRAMDEYTSSEIFMGKSTIYLAHNTC	420	
QY	421	EDSLAAPILIDLVL	LAELSTRIEFKAENEGK	FHSFHPVATILSYLTAKALVP	PCTPVVN	480
DB	421	EDSLAAPILIDLVL	LAELSTRIOFKAENEGK	FHSFHPVATILSYLTAKALVP	PCTPVVN	480
QY	481	ALSKORAMLENIMRACV	GLAPENNMILEYK	510		
DB	481	ALSKORAMLENIMRACV	GLAPENNMILEYK	510		

RESULT 11
US-10-718-952-16
;; Sequence 16, Application US/10718952
;; Publication No. US20040128713A1
;; GENERAL INFORMATION:
;; APPLICANT: Hitz, William
;; APPLICANT: Sebastian, Scott
;; APPLICANT: Grace, John
;; APPLICANT: Streitz, Leon
;; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
;; FILE REFERENCE: BB-1077-C
;; CURRENT APPLICATION NUMBER: US/10/718,952
;; PRIOR FILING DATE: 2003-11-21
;; PRIOR APPLICATION NUMBER: 08/835,751
;; PRIOR FILING DATE: APRIL 8, 1997
;; PRIOR APPLICATION NUMBER: PCT/US98/06822
;; PRIOR FILING DATE: APRIL 7, 1998
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 16

;; LENGTH: 510
;; TYPE: PRT
;; ORGANISM: Glycine max
US-10-718-952-16

Query Match 98.8%; Score 2600; DB 16; Length 510;
Best Local Similarity 98.6%; Pred. No. 5.2e-217;
Matches 503; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY	1	MEIENFKVCEPNVKTETETI	QSVNYVNETTEL	VHENRNGTYQMI	VKPKSVKEFKNTINHP	60
DB	1	MEIENFKVCEPNVKTETETI	QSVNYVNETTEL	VHENRNGTYQMI	VKPKSVNYQFKNTINHP	60
QY	61	KLGVMLVGMGNGNSTLTG	VIANREGISWATKDKI	QOANYFGSLTQASAIRVGS	FOGEE	120
DB	61	KLGVMLVGMGNGNSTLTG	VIANREGISWATKDKI	QOANYFGSLTQASAIRVGS	FOGEE	120
QY	121	IYAPFKSLIPMNPPD	IVFGGWDISNMNLADMA	RAKVFDDLOKOLRPYMESMLPLPGI	180	
DB	121	IYAPFKSLIPMNPPD	IVFGGWDISNMNLADMA	RAKVFDDLOKOLRPYMESMLPLPGI	180	
QY	181	YDPDFIAANOEERANNVI	KGTKOEVOQI	IKDIAKFKATKVDKVVVLTANTERYSNLV	240	
DB	181	YDPDFIAANOEERANNVI	KGTKOEVOQI	IKDIAKFKATKVDKVVVLTANTERYSNLV	240	
QY	241	VGNDTMENLLAADVDRNEA	ISPSSTLYAIACWENVP	FINSGPONTFVPGILDLAIANT	300	
DB	241	VGNDTMENLLAADVDRNEA	ISPSSTLYAIACWENVP	FINSGPONTFVPGILDLAIANT	300	
QY	301	LIGGDDFKSGQTKMSV	VDPLVAGIKPTSTIVSYNH	GNNDGMLSAPOFRSKEISKS	360	
DB	301	LIGGDDFKSGQTKMSV	VDPLVAGIKPTSTIVSYNH	GNNDGMLSAPOFRSKEISKS	360	
QY	361	NVDDMVNSNALIYEPGE	HPHVVIKYVPYVGS	KRAMDEYTSSEIFMGKSTIYLAHNTC	420	
DB	361	NVDDMVNSNALIYEPGE	HPHVVIKYVPYVGS	KRAMDEYTSSEIFMGKSTIYLAHNTC	420	
QY	421	EDSLAAPILIDLVL	LAELSTRIEFKAENEGK	FHSFHPVATILSYLTAKALVP	PCTPVVN	480
DB	421	EDSLAAPILIDLVL	LAELSTRIOFKAENEGK	FHSFHPVATILSYLTAKALVP	PCTPVVN	480
QY	481	ALSKORAMLENIMRACV	GLAPENNMILEYK	510		
DB	481	ALSKORAMLENIMRACV	GLAPENNMILEYK	510		

RESULT 12
US-10-025-003-12
;; Sequence 12, Application US/10025003
;; Publication No. US20030074685A1
;; GENERAL INFORMATION:
;; APPLICANT: Hitz, William
;; APPLICANT: Sebastian, Scott
;; APPLICANT: Grace, John
;; APPLICANT: Streitz, Leon
;; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
;; FILE REFERENCE: BB-1077-C
;; CURRENT APPLICATION NUMBER: US/10/025,003
;; PRIOR FILING DATE: 2002-05-07
;; PRIOR APPLICATION NUMBER: 08/835,751
;; PRIOR FILING DATE: APRIL 8, 1997
;; PRIOR APPLICATION NUMBER: PCT/US98/06822
;; PRIOR FILING DATE: APRIL 7, 1998
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 12
;; LENGTH: 510
;; TYPE: PRT
;; ORGANISM: Glycine max
US-10-025-003-12

Query Match 98.5%; Score 2593; DB 14; Length 510;

Best Local Similarity 98.4%; Pred. No. 2.1e-216;
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MEIENFKVCEPVNKTETETIOSVYVETTELVEHNRNGTYQWIVKPKSVKTEFKTNHVP 60
DB 1 MEIENFKVCEPVNKTETETIOSVYVETTELVEHNRNGTYQWIVKPKSVKTEFKTNHVP 60
QY 61 KLGWLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
DB 61 KLGWLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120
QY 121 IYAPFKSLPWNPPDIVEFGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
DB 121 IYAPFKSLPWNPPDIVEFGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180
QY 181 YDPDFIAANOEBERANVVKGTQOEVOQIIKDIKAFKATKVDKVVLTANTATERSNLV 240
DB 181 YDPDFIAANOEBERANVVKGTQOEVOQIIKDIKAFKATKVDKVVLTANTATERSNLV 240
QY 241 VGLNDTMENTLAAVDNRNEAISPSTLYAIACMENVPTNGSPONTFVPGGLIDLAIAART 300
DB 241 VGLNDTMENTLAAVDNRNEAISPSTLYAIACMENVPTNGSPONTFVPGGLIDLAIAART 300
QY 301 LIGDDPFSGQTKMSVLVDPLVAGIKPTSIYSYNHLGNDGMNLSAPOTRSKEISKS 360
DB 301 LIGDDPFSGQTKMSVLVDPLVAGIKPTSIYSYNHLGNDGMNLSAPOTRSKEISKS 360
QY 361 NVDDMVNSNALIYEPGEHPDHVVVVKVPPYVDSKRAMDEYTSBIFMGKSTIVLHNTC 420
DB 361 NVDDMVNSNALIYEPGEHPDHVVVVKVPPYVDSKRAMDEYTSBIFMGKSTIVLHNTC 420
QY 421 EDSLAAPIILDVLVLAELSTRIEPKAENEGKFSHPVATILSYLTAKPLVPGTPPVN 480
DB 421 EDSLAAPIILDVLVLAELSTRIEPKAENEGKFSHPVATILSYLTAKPLVPGTPPVN 480
QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 13

US-10-718-952-12
; Sequence 12, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718, 952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835, 751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
US-10-718-952-12

Query Match 98.5%; Score 2593; DB 16; Length 510;
Best Local Similarity 98.4%; Pred. No. 2.1e-216;
Matches 502; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MEIENFKVCEPVNKTETETIOSVYVETTELVEHNRNGTYQWIVKPKSVKTEFKTNHVP 60
DB 1 MEIENFKVCEPVNKTETETIOSVYVETTELVEHNRNGTYQWIVKPKSVKTEFKTNHVP 60

QY 61 KLGWLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120

DB 61 KLGWLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120

QY 121 IYAPFKSLPWNPPDIVEFGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180

DB 121 IYAPFKSLPWNPPDIVEFGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180

QY 181 YDPDFIAANOEBERANVVKGTQOEVOQIIKDIKAFKATKVDKVVLTANTATERSNLV 240

DB 181 YDPDFIAANOEBERANVVKGTQOEVOQIIKDIKAFKATKVDKVVLTANTATERSNLV 240

QY 241 VGLNDTMENTLAAVDNRNEAISPSTLYAIACMENVPTNGSPONTFVPGGLIDLAIAART 300

DB 241 VGLNDTMENTLAAVDNRNEAISPSTLYAIACMENVPTNGSPONTFVPGGLIDLAIAART 300

QY 301 LIGDDPFSGQTKMSVLVDPLVAGIKPTSIYSYNHLGNDGMNLSAPOTRSKEISKS 360

DB 301 LIGDDPFSGQTKMSVLVDPLVAGIKPTSIYSYNHLGNDGMNLSAPOTRSKEISKS 360

QY 361 NVDDMVNSNALIYEPGEHPDHVVVVKVPPYVDSKRAMDEYTSBIFMGKSTIVLHNTC 420

DB 361 NVDDMVNSNALIYEPGEHPDHVVVVKVPPYVDSKRAMDEYTSBIFMGKSTIVLHNTC 420

QY 421 EDSLAAPIILDVLVLAELSTRIEPKAENEGKFSHPVATILSYLTAKPLVPGTPPVN 480

DB 421 EDSLAAPIILDVLVLAELSTRIEPKAENEGKFSHPVATILSYLTAKPLVPGTPPVN 480

QY 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510
DB 481 ALSKORAMLENIMRACVGLAPENNMILEYK 510

RESULT 14

US-10-424-599-154864
; Sequence 154864, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424, 599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154864
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110863C.1.pep
US-10-424-599-154864

Query Match 95.9%; Score 2524; DB 15; Length 510;
Best Local Similarity 94.7%; Pred. No. 2.1e-210;
Matches 483; Conservative 18; Mismatches 9; Indels 0; Gaps 0;

QY 1 MEIENFKVCEPVNKTETETIOSVYVETTELVEHNRNGTYQWIVKPKSVKTEFKTNHVP 60

DB 1 MEIENFKVCEPVNKTETETIOSVYVETTELVEHNRNGTYQWIVKPKSVKTEFKTNHVP 60

QY 61 KLGWLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120

DB 61 KLGWLVGMGNGNSTLTGVIANREGISWATKDKIQOANYFGSLTQASAIRVGSFOGEE 120

QY 121 IYAPFKSLPWNPPDIVEFGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180

DB 121 IYAPFKSLPWNPPDIVEFGWDISNMNLADAMARAKVFDIDLOKOLRPYMESMLPLPGI 180

QY 181 YDPDEIAANOEBRANNVIKGTQEOVOQI IKD IKAFKATKVYKVYVMTANTERYSNLV 240
DB 181 YDPDEIAANOGBRATNVIRGTRKKEQVOQI IKD IREFKSTKVYKVYVMTANTERYSNLV 240
QY 241 VGLNNTMELLAAVDRNEAEISPSLTIXAIACVMENVPFINSPOPTFVPGGLIDLAIAKNT 300
DB 241 VGLNNTMELLAAVDRNEAEISPSLTIXAIACVMENVPFINSPOPTFVPGGLIDLAIAKNT 300
QY 301 LIGGDFKSGQTKMSVLVDFLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360
DB 301 LIGGDFKSGQTKMSVLVDFLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS 360
QY 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGKSTIVLAHNTC 420
DB 361 NVDDMVNSNALIYEPGEHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGKSTIVLAHNTC 420
QY 421 EBSLAAPIILDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPVNVNALSQ 480
DB 421 EBSLAAPIILDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPVNVNALSQ 480
QY 481 ALSKORAMLENTIRACVGLAPENNMLEYK 510
DB 481 ALSKORAMLENTIRACVGLAPENNMLEYK 510

RESULT 15
US-10-442-017-15

; Sequence 15, Application US/10442017
; Publication No. US20030215950A1
; GENERAL INFORMATION:
; APPLICANT: Battelle Memorial Institute
; APPLICANT: Lasure, Linda L.
; APPLICANT: Dai, Ziyu
; TITLE OF INVENTION: Isolated Polynucleotides and Methods of Promoting a Morphology In
; TITLE OF INVENTION: a Fungus
; FILE REFERENCE: BAA-195
; CURRENT APPLICATION NUMBER: US/10/442,017
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: 60/382,132
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 15
; LENGTH: 505
; TYPE: PRF
; ORGANISM: Sesamum indicum
US-10-442-017-15

Query Match 92.4%; Score 2432; DB 15; Length 505;

Best Local Similarity 92.3%; Pred. No. 2,1e-202;

Matches 466; Conservative 19; Mismatches 20; Indels 0; Gaps 0;

QY 6 FVECEPNVYKTTETELOSVYVNETTELHNRNGTQWIVPKSVKYEFTKTIHVPKLGVM 65
DB 1 FVECEPNVYKTTETELHSHVYVNETTELHNRNGTQWIVPKSVKYEFTKTIHVPKLGVM 60
QY 66 LVGMGNNGSTLTGVIANREGISWATKDIQOANYFGSLTOASAIRVGSFOGEBIYAPF 125
DB 61 LVGMGNNGSTLTGVIANREGISWATKDIQOANYFGSLTOASAIRVGSFOGEBIYAPF 120
QY 126 KSLLPVNDPDIIVFGGWDISNNMLADAMARAKVFDIDLQKLRPYMESMLPLGIYDPDF 185
DB 121 KSLLPVNDPDIIVFGGWDISNNMLADAMARAKVFDIDLQKLRPYMEHVPPLGIYDPDF 180
QY 186 IAAOGEERANNVIKGTQEOVOQI IKD IKAFKATKVYKVYVMTANTERYSNLVGND 245
DB 181 IAAOGEERANNVIKGTQEOVOQI IKD IKAFKATKVYKVYVMTANTERYSNLVGND 240
QY 246 TWENTLLAAVDRNEAEISPSLTIXAIACVMENVPFINSPOPTFVPGGLIDLAIAKNTLIGD 305
DB 241 TWENTLLAAVDRNEAEISPSLTIXAIACVMENVPFINSPOPTFVPGGLIDLAIAKNTLIGD 300
QY 306 DFSGQTKMSVLVDFLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS NVDD 365

DB 301 DFSGQTKMSVLVDFLVGAGIKPTSIYSYNHLGNDGMNLSAPQTFRSKEISKS NVDD 360
QY 366 MWSNAILYEPGEHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGKSTIVLAHNTCEDSL 425
DB 361 MWSNAILYEPGEHPDHVVVIKYVPYVDSKRAMDEYTSSEIFMGKSTIVLAHNTCEDSL 420
QY 426 AAPILDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPVNVNALSQ 485
DB 421 AAPILDLVLAELSTRIEFKAENEGKFSFHPVATILSYLTKAPLVPGTPVNVNALSQ 480
QY 486 RAMLENTIRACVGLAPENNMLEYK 510
DB 481 RAMLENTIRACVGLAPENNMLEYK 505

Search completed: June 7, 2005, 17:03:07
Job time : 84.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model1

Run on: June 7, 2005, 16:31:27 ; Search time 31 Seconds
(without alignments)
1228.098 Million cell updates/sec

Title: US-10-718-952-2
Perfect score: 2632
Sequence: 1 MFIENFKVCEPVNXYTETETI.....NIMRACVGLAPENNILEYK 510

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/6CTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile01.pep:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2336	88.8	510	3	US-09-118-442-11 Sequence 11, Appl
2	2336	88.8	510	3	US-09-677-064-11 Sequence 11, Appl
3	2335	88.7	510	4	US-09-727-628-2 Sequence 2, Appl
4	1382.5	52.5	533	4	US-09-734-237B-73 Sequence 73, Appl
5	1382.5	52.5	534	4	US-09-734-237B-75 Sequence 75, Appl
6	1376.5	52.3	525	4	US-09-248-796A-17234 Sequence 17234, A
7	541	20.6	444	4	US-09-902-540-12518 Sequence 12518, A
8	113.5	4.3	730	4	US-09-107-433-4707 Sequence 4707, Ap
9	113.5	4.3	731	4	US-09-583-110-4720 Sequence 4720, Ap
10	112.5	4.2	452	4	US-09-489-039A-11570 Sequence 11570, A
11	111.5	4.2	452	4	US-09-248-796A-20645 Sequence 20645, A
12	111.5	4.2	731	1	US-09-014-897-2 Sequence 2, Appl
13	111.5	4.2	731	1	US-08-731-716-2 Sequence 2, Appl
14	111	4.2	935	4	US-09-512-250C-33 Sequence 33, Appl
15	109.5	4.2	747	4	US-09-583-110-3066 Sequence 3066, Ap
16	109.5	4.2	751	4	US-09-107-433-3375 Sequence 3375, Ap
17	108	4.1	2465	2	US-08-596-291-3 Sequence 3, Appl
18	108	4.1	2465	3	US-09-100-804-3 Sequence 3, Appl
19	108	4.1	2466	3	US-09-080-855-12 Sequence 12, Appl
20	108	4.1	2466	4	US-09-566-076-12 Sequence 12, Appl
21	108	4.1	2466	5	PCT-US94-09943-2 Sequence 2, Appl
22	108	4.1	2485	3	US-09-290-640-46 Sequence 46, Appl
23	108	4.1	2485	4	US-09-665-615B-46 Sequence 46, Appl
24	107.5	4.1	1963	4	US-09-583-110-5243 Sequence 5243, Ap
25	107.5	4.1	1967	4	US-09-107-433-4883 Sequence 4883, Ap
26	106.5	4.0	778	4	US-09-489-039A-8455 Sequence 8455, Ap
27	106	4.0	539	4	US-09-543-681A-6877 Sequence 6877, Ap

28	106	4.0	1242	4	US-09-107-532A-5241 Sequence 5241, Ap
29	103.5	3.9	912	4	US-09-248-796A-20032 Sequence 20032, A
30	101.5	3.9	2042	4	US-09-077-098A-6 Sequence 6, Appl
31	101	3.8	449	4	US-09-107-433-4132 Sequence 4132, Ap
32	101	3.8	454	4	US-09-583-110-3922 Sequence 3922, Ap
33	100.5	3.8	545	4	US-09-248-796A-14131 Sequence 14131, A
34	100.5	3.8	835	4	US-09-438-185A-728 Sequence 728, Ap
35	100.5	3.8	10182	3	US-09-134-001C-3159 Sequence 3159, Ap
36	99.5	3.8	390	3	US-08-745-995A-18 Sequence 18, Appl
37	99.5	3.8	390	4	US-09-005-352-18 Sequence 18, Appl
38	99.5	3.8	420	3	US-08-745-995A-14 Sequence 14, Appl
39	99.5	3.8	420	4	US-09-005-352-14 Sequence 14, Appl
40	99.5	3.8	1091	4	US-09-583-110-3823 Sequence 3823, Ap
41	99.5	3.8	1102	4	US-09-107-433-4859 Sequence 4859, Ap
42	99	3.8	108	4	US-09-489-039A-13228 Sequence 13228, A
43	99	3.8	414	4	US-09-107-433-3327 Sequence 3327, Ap
44	99	3.8	463	4	US-09-710-279-960 Sequence 960, App
45	99	3.8	551	4	US-09-489-039A-11049 Sequence 11049, A

ALIGNMENTS

```

RESULT 1
US-09-118-442-11
; Sequence 11, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Faelsrq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-118-442-11
Query Match      88.8%; Score 2336; DB 3; Length 510;
Best Local Similarity 87.3%; Pred. No. 4.5e-223;
Matches 445; Conservative 31; Mismatches 34; Indels 0; Gaps 0;
1 MFIENFKVCEPVNXYTETETIQLVHNRNGTYOIVKPKSVYKFKNIHVP 60
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
1 WPIESFRVSPHVRGPMIEBSERYDTELVEHGDSRVRVBPVKVFNFRITAVP 60
KLGWLVGCGNGGSLTLAGVIANREGISWATKDVQOANVYGSILTHASTIVGSNGER 120
61 KLGWLVGCGNGGSLTLAGVIANREGISWATKDVQOANVYGSILTHASTIVGSNGER 120
61 KLGWLVGCGNGGSLTLAGVIANREGISWATKDVQOANVYGSILTHASTIVGSNGER 120
121 IYAPFKSLIPVNPDIYFGGWDINMNLADAMARAKFDIDLOQLRPYMESMLPLPGI 180
121 IYAPFKSLIPVNPDIYFGGWDINMNLADAMARAKFDIDLOQLRPYMESMLPLPGI 180
121 IYAPFKSLIPVNPDIYFGGWDINMNLADAMARAKFDIDLOQLRPYMESMLPLPGI 180
181 YDPDFIANQBERANVIGTKQEOVOQIINDIKAFKATKVKVAVVMTANTERYSNLV 240
181 YDPDFIANQBERANVIGTKQEOVOQIINDIKAFKATKVKVAVVMTANTERYSNLV 240
181 YDPDFIANQBERANVIGTKQEOVOQIINDIKAFKATKVKVAVVMTANTERYSNLV 240
181 YDPDFIANQBERANVIGTKQEOVOQIINDIKAFKATKVKVAVVMTANTERYSNLV 240
241 VGLNPTMNLAAVDRNBAEISPTLYAIACMENVPPINGSQPTFPVGLIDLAIARNT 300

```

```
Db 241 AGANDTMEVLASVDKNEAEVSPSTLYAIAACMEGVFINSPOPTFVPGILDIAIKNNC 300
Qy 301 LIGGDDFSGQTKMKSVLVDPLVGAIGIKPTISIVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Db 301 LIGGDDFSGQTKMKSVLVDPLVGAIGIKPTISIVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALILYEPGHPHVVVVKVYPVYVGSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVDDMVNSNALILYEPGHPHVVVVKVYPVYVGSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAEISTRIFEKAEENKGFHSPHVATILSYLTKAFLVPPGTPVNV 480
Db 421 EDSLAAPIIIDLVLLAEISTRIFEKAEENKGFHSPHVATILSYLTKAFLVPPGTPVNV 480
Qy 481 ALSQRAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALSQRAMLENIMRACVGLAPENNMLEYK 510
```

RESULT 2

```
US-09-677-064-11
; Sequence 11, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-677-064-11
```

```
Query Match 88.8%; Score 2336; DB 3; Length 510;
Best Local Similarity 87.3%; Pred. No. 4,5e-223;
Matches 445; Conservative 31; Mismatches 34; Indels 0; Gaps 0;
```

```
Qy 1 MFIEFKVCEPNVKTETEISGVNYETTELVHNRNGTYOMIVKPKSVKYEFTKNIHVP 60
Db 1 MFIEFVRVSPHVRGPEIESEYRDTTELVEHKGDSRWVVRPKSVKYNFRRTAVP 60
Qy 61 KLGWVLVGMGNGNSTLTGVIYANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGE 120
Db 61 KLGWVLVGMGNGNSTLTGVIYANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGE 120
Qy 121 IYAPFKSLPLPMVNPDDIVFGMDISNMNLADAMARAKYFDIDLOKOLRPYMESMLPLPGI 180
Db 121 IYAPFKSLPLPMVNPDDIVFGMDISNMNLADAMARAKYFDIDLOKOLRPYMESMLPLPGI 180
Qy 181 YDPDFIAANQSGRANNVKGTKEQVEQIIOIKRFEKATKVDKVVVMTANTERYSNLV 240
Db 181 YDPDFIAANQSGRANNVKGTKEQVEQIIOIKRFEKATKVDKVVVMTANTERYSNLV 240
Qy 241 VGLNDTMEVLAAVDNRNEAISPSTLYAIAACMEVNPVINSPOPTFVPGILDIAIAKNT 300
Db 241 VGLNDTMEVLAAVDNRNEAISPSTLYAIAACMEVNPVINSPOPTFVPGILDIAIAKNT 300
Qy 361 NVDDMVNSNALILYEPGHPHVVVVKVYPVYVGSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVDDMVNSNALILYEPGHPHVVVVKVYPVYVGSKRAMDEYTSSEIFMGKSTIVLHNTC 420
```

```
Qy 301 LIGGDDFSGQTKMKSVLVDPLVGAIGIKPTISIVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Db 301 LIGGDDFSGQTKMKSVLVDPLVGAIGIKPTISIVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALILYEPGHPHVVVVKVYPVYVGSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVDDMVNSNALILYEPGHPHVVVVKVYPVYVGSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Qy 421 EDSLAAPIIIDLVLLAEISTRIFEKAEENKGFHSPHVATILSYLTKAFLVPPGTPVNV 480
Db 421 EDSLAAPIIIDLVLLAEISTRIFEKAEENKGFHSPHVATILSYLTKAFLVPPGTPVNV 480
Qy 481 ALSQRAMLENIMRACVGLAPENNMLEYK 510
Db 481 ALSQRAMLENIMRACVGLAPENNMLEYK 510
```

RESULT 3

```
US-09-727-628-2
; Sequence 2, Application US/09727628
; Patent No. 6791013
; GENERAL INFORMATION:
; APPLICANT: Armstrong, Katherine
; APPLICANT: Hey, Timothy D
; APPLICANT: Folkerts, Otto
; APPLICANT: Smith, Kelley A
; APPLICANT: Hopkins, Nicole L
; TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
; FILE REFERENCE: 50597
; CURRENT APPLICATION NUMBER: US/09/727,628
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 60/168,612
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Zea mays
US-09-727-628-2
```

```
Query Match 88.7%; Score 2335; DB 4; Length 510;
Best Local Similarity 87.6%; Pred. No. 5.6e-223;
Matches 447; Conservative 29; Mismatches 34; Indels 0; Gaps 0;
```

```
Qy 1 MFIEFKVCEPNVKTETEISGVNYETTELVHNRNGTYOMIVKPKSVKYEFTKNIHVP 60
Db 1 MFIEFVRVSPHVRGPEIESEYRDTTELVEHKGDSRWVVRPKSVKYNFRRTAVP 60
Qy 61 KLGWVLVGMGNGNSTLTGVIYANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGE 120
Db 61 KLGWVLVGMGNGNSTLTGVIYANREGISWATKDKIQOANYFGSLTQASAIRVSGFQGE 120
Qy 121 IYAPFKSLPLPMVNPDDIVFGMDISNMNLADAMARAKYFDIDLOKOLRPYMESMLPLPGI 180
Db 121 IYAPFKSLPLPMVNPDDIVFGMDISNMNLADAMARAKYFDIDLOKOLRPYMESMLPLPGI 180
Qy 181 YDPDFIAANQSGRANNVKGTKEQVEQIIOIKRFEKATKVDKVVVMTANTERYSNLV 240
Db 181 YDPDFIAANQSGRANNVKGTKEQVEQIIOIKRFEKATKVDKVVVMTANTERYSNLV 240
Qy 241 VGLNDTMEVLAAVDNRNEAISPSTLYAIAACMEVNPVINSPOPTFVPGILDIAIAKNT 300
Db 241 VGLNDTMEVLAAVDNRNEAISPSTLYAIAACMEVNPVINSPOPTFVPGILDIAIAKNT 300
Qy 301 LIGGDDFSGQTKMKSVLVDPLVGAIGIKPTISIVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Db 301 LIGGDDFSGQTKMKSVLVDPLVGAIGIKPTISIVSYNHLGNNDGNMLSAPQTFRSKEISKS 360
Qy 361 NVDDMVNSNALILYEPGHPHVVVVKVYPVYVGSKRAMDEYTSSEIFMGKSTIVLHNTC 420
Db 361 NVDDMVNSNALILYEPGHPHVVVVKVYPVYVGSKRAMDEYTSSEIFMGKSTIVLHNTC 420
```

OY	421	EDSLAAPIIDLVLAELSTRIFPAEKEGFSPHPVATILSYTTKPVPVPGPPVN	480
Dd	421	EDSLAARIIDVLVAELSTRIQLKPEBDTRFHSHPVATILSYTTKPVPVPGPPVN	480
OY	481	ALSKORALENIMRACVGAPENNMILEXX	510
Dd	481	ALAKORALENIMRACVGAPENNMLEXX	510

```

RESULT 4
US-09-734-237B-73
: Sequence 73. Application US/09734237B
: Patent No. 6618752
: GENERAL INFORMATION:
: APPLICANT: Kozzelli, J. David
: APPLICANT: Bu, Peter
: APPLICANT: Hua, Ling
: TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
: FILE REFERENCE: 5853.10608
: CURRENT APPLICATION NUMBER: US/09/734,237B
: PRIORITY FILING DATE: 2000-12-08
: PRIOR APPLICATION NUMBER: 09/494,921
: NUMBER OF SEQ ID NOS: 79
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 73
: LENGTH: 533
: TYPE: prt
: ORGANISM: Saccharomyces cerevisiae
: US-09-734-237B-73

```

Query Match	52.5%	Score 1382.5;	DB 4;	Length 533;
Best Local Similarity	52.9%;	Pred. No. 2.7e-128;		
Matches 277;	Conservative 86;	Mismatches 140;	Indels 21;	Gaps 8;

QY	3	IEHFKCEPVRKXTEHIEISVYNYETTELHYENRNGTQWVLYKPSVXKEEKTNIHP-K	61
Dp	9	ITSVKKVTDKCYKDNELLTKISYEBA-VVTKASGRF--VTPPVODVFKDLKKREK	65
QY	62	LGVMVLWGCGNNGSTLTGVIANREGISWATYDKIQQANYPGSLTQSAIRVG-SFOGE	120
Dp	66	LGIMLLIGLGGNGSTLVAIVLANKHVEPQTKEGVKQPNYFCSMTQSTLKLGIADGND	125
QY	121	IYAPFSLLPYNPPEDIVFGWDISNMMLADMAAKYFDDLDQOLAPRYESMLPLPGI	180
Dp	126	VYAPFSLPLMWSBNDPVVSGWDINNADLEYMOSQVLEVDLQORLAKKSLVKPPESI	185
QY	181	YDPDFIAANOERANNVI---KGT-----KQOVQOQIKDKAKFEATKVDKVVLMTA	231
Dp	186	YYDPFIAANODERANNVCINLDEKGNVTRGKWTHTQIRLRDIOQNKEMALDKYIVLMTA	245
QY	232	NTERYVENLVGNDTMENTLAAVDNRNEAISTLYALACMENVPFNSPQNTFVPGI	291
Dp	246	NTERYVEVBPVNDTMENTLQSIKDNHEIASTTFAAASILEGVPYINGSPQNTFVPGI	305
QY	292	IDLAIRANTLLIGDDPKSGQTPMKSVLYDVLGAGIKFSTISVYNHGLGNDGNMLSAQOT	351
Dp	306	VOLAHGEGTFIAGDDPKSGQTKLSVLAQFLVDAIKFVSIASVYNHGLGNDGNMLSAQKQ	365
QY	352	FRSEKESKNVVDMDWNSNAIILYEP--GEHPHVVIKYVYVYDGSKRANDEYSLEFMG	409
Dp	366	FRSEKESKSSVDDIITIASNDILYNKDLGKVDHCIVITKMPVDDSKVADEYVSEIMLG	425
QY	410	GKSTVLAHNTCEDSLAPRIILDVLVLAELSTRIEFK-----AENEGKESFHHVATILS	464
Dp	426	GHNRIISIHVCEDSLATPLRIIDLVMTHEPCFRVSYKKVDPVKDAGKGFENFVVLTFLS	485
QY	465	YLTKAPLVEPPTPVYVALSKQAPAMENIMRACVGLAPENNMILE	508
Dp	486	YMLKAPLTPGPHVPGVNGLKORTALENFRLLIGPSONEIRFR	529

```

US-09-734-237B-75
; Sequence 75, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bul, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583/40608
; CURRENT APPLICATION NUMBER: US/09/734,237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494,921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic protein derived from Saccharomyces cerevisiae myo-inositol
; OTHER INFORMATION: tol-1-phosphate synthase, having a glycine residue inserted after
; OTHER INFORMATION: the initiating methionine
US-09-734-237B-75

```

Query Match	52.5%;	Score 1382.5;	DB 4;	Length 534;
Best Local Similarity	52.9%;	Pred. No. 2.7e-128;		
Matches 277;	Conservative 86;	Mismatches 140;	Indels 21;	Gaps 8;

Qy	1	IENFKCEPJVKTETHEISVYVNETELVHENRGTQWIVKPSVYCEKTNHNP-K	61
Dh	10	ITSVKVTDKCTYKONELLTKSYENA-VTKTASGRD--VTPVVDYVKLDKKEPK	66
Qy	62	LGWLVGCGNGNSTLTGCVIANREGISWATKDKIQANYFGSLTQSAIRVG-SFOGEE	120
Dh	67	LGIMLLGLGNGNSTLTVASVLANKNVEFQTEKEGVQENYFGSMTCSTLLXGIDAGND	126
Qy	121	IYAPFSLLPYNPDDIVFGWMDISNNMLADMAAKFDDILOKOLRPYMESMLPLPGI	180
Dh	127	VYAPFSLPLPMSPNDPVGVSMDINNADYEMQBSVLETDLOORLAKMSLVKPLPSI	186
Qy	181	YDEDFIAANQEBRANVVI-----KGT-----KQEQVOQIIKDKIKAFKATKVDKVVLMTA	231
Dh	187	YYEDFIAANQEBRANNCINLDEKGVNTRGKWNHQRIRRDIOQPKENALDKIVLMTA	246
Qy	232	NTERYSNLVYGNDPMENLLAVDNEAEMISPTIYALACMEVNPPTNGSPONTYPGL	291
Dh	247	NTERYSNPVGNDPMENLLQSIKNDHEILASTIIPAASILLEGPPYINGSPONTYPGL	306
Qy	292	IDLAIRANLTIIGDDPFKSGQTKMKSVLVDPLVAGAIKPTSIYSVYHGLGNDGNMLISAPQT	351
Dh	307	VQLAHEGFTIAGDDPLKSGQTKLKSVLQPLVDAGIKFVSIASVNHGNDGYNLSPKQ	366
Qy	352	FRKEKISKNNVDDMWSNNAIUYEE--GEHDPVVYIKYVYVYDSDSKRANDEYSELFMG	408
Dh	367	FRKEKISKSSVIDDIIASNDILYNDKLNKVDHCIVITMKPVDSDSKRANDEYSELMLG	428
Qy	410	GKSTVLAHNTCEDSLIAPILIDVLALAELESTRIEFK-----AENBEKFSFHPVATILS	464
Dh	427	GHNRIISIHVCEDSLATPLTIDLLVMTEFCRVSYKKVDPVKDAKGFENFYVLTFLS	486
Qy	465	YLTKAPLVPPTPVYNALSKQAPMLNEMRACVGLAPENNMILE508	
Dh	487	YMLKAPLTPGFHPVNGLNKQTALENLRLLLIGLPSQENELRFE530	

Qy	465	YLTRAPLVPGSTPVYNALSKORAMLENIMRACVGLAPENNWILE	508
Db	486	YMLKAPLTPRPGFHHVNGINKORTALENLRLLILPSQNEIRFE	529

RESULT 5

RESULT 6

US-09-248-796A-17234

; Sequence 17234, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

```

; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17234
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17234

Query Match      52.3%; Score 1376.5; DB 4; Length 525;
Best Local Similarity 52.2%; Pred. No. 1e-127;
Matches 263; Conservative 98; Mismatches 128; Indels 15; Gaps 5;

QY 17 ETEISVNYETTELIVHENRNGTYOMIVKPKSVKYEFTKINIHVPKLGVMVGMGNGNST 76
DB 21 DDLTKTKTYENS-VKEDANG--KFIVTPASDYEFKVDLKVPRKVGILLVIGGNGT 77
QY 77 LTGVIANREGISWATKQIQOANYFGSLTQASAIRVG--SFQGEIYAPFKSLIPVNP 134
DB 78 LLGATLADKHNISFENKEGVKPNYGSVYQASIVKIGVDKETGBDVYVFPFNSIVPMVP 137
QY 135 DDIVGCGDISNMNLADMAKAVPDIDLQKQREPMSMLPLPGIYDDPFIANAQERA 194
DB 138 NDLVGWDISGLPLDQAKRAKVLDTVLQKQLYPLEKKPLLESITYPDFIALNQSEBA 197
QY 195 NNVIK-----GTKQBOVOQIIKDIAKFKATKVDKVVLMNTANTERSMLVVGNDT 246
DB 198 NNVPQVNGEVKTDKMAKADVEKIRDIRPKAKNELDKYIIIMTANTREKAVLVNVDI 257
QY 247 MENLLAADRNEAESIPSTLYAIACMVENVPFINSPONTFVPLGLIDLAIAANTLIGDD 306
DB 258 ADNLKISIKESHEELAPSTVFASILEKVPYINSPOQTFVPVGYELAEKYDSFEGDD 317
QY 307 FRSQGTAKSVYVDPLVNGIGIKPTISVYNHGNNDGMLSPQFRSKSEIKSNVNDM 366
DB 318 FRSQGTAKSVYAOPLVNDAGIKPLSLASYNHGNNDGMLSPKFRSKSEIKSQVVDI 377
QY 367 VNSNALY--EPGEHPDHVVIVKVPYVDSKRAMDEYSEIFMGSKSTIVLHNTCEDSL 424
DB 378 IESNELYNKESGDVDCIVIKIIPAYDSKVADEYSELMGHNKSIHNCEBSL 437
QY 425 LAAPITLDVLLAELSTRIEFAENEGKFSHPVATITISYTKAPLVPPGTPVNVALS 484
DB 438 LATPLILDVVAETFAIRVQVKGPKSDYDELVPVASTLYMLKAPLAPGFPINGLAK 497
QY 485 GRAMLENIRACVGLAPENNMILE 508
DB 498 ORQQLVNLISLVGLPIDNELRFE 521

RESULT 7
US-09-902-540-12518
; Sequence 12518, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12518
```

```

; LENGTH: 444
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12518

Query Match      20.6%; Score 541; DB 4; Length 444;
Best Local Similarity 29.8%; Pred. No. 8.3e-45;
Matches 140; Conservative 92; Mismatches 190; Indels 48; Gaps 12;

QY 52 EFKTNIHVP--KLGMVLVGMGNGNSTLTGVIANREGISWATKQIQOANYFGSLTQAS 109
DB 2 ENKSVAKPDGLAVLILGLAVS--TILMAGVELARKG-----KHPISGLTQMG 50
QY 110 AIRVGS--FQGEIYAPFKSLIPVNPDDIVEGCGDISNMNLADMAKAVPDIDLQKQLR 168
DB 51 TARIKRTDGRTV--KINELVPLAEIKQVAFGAWDIIREDAYEVAVRGSLDKILEVK 108
QY 169 PYMSMLPLPGIYDDPFIANAQERANVIGTK--QGVQOQIIKDIAKFKATKVDKVP 226
DB 109 PPLQSIKPKQGVHDEFEFRIRIEA---NHKATKTHRESIALRQDIRDFKELNATRAV 164
QY 227 VLMNTANTERSNVLVVGNDTMENTLLAADRNEAESIPSTLYAIACMVENVPFINSPO-N 285
DB 165 MVVSSVETFPRLPESFK--TLAAREKALDENSEPDINPTALTYAIKGVFPANATPNA 223
QY 286 TFPVPLIDLAIAIRNTLIGDDPKSGQTKMSVLVDPLVAGIKPTISVYNHGNNDGMN 345
DB 224 VDPALQEMAQESVAVAGRLKSGQTMKTIVAPLAKMLGLDGMFSTNILGRDGEV 283
QY 346 LSAPOFRSKSEIKSNVNDVNVNSALILEPGEHPD-----HYVVIKYVYVGSKRAM 399
DB 284 LDDPAFAFAKAVTKSSVVDITL-----QPDLYBELKYYAKHVAHIYYPRGDAKSGW 336
QY 400 DEYSEIMGGSSTVLHNTCEDSLAAPTITLDVLTLAELSTRIEPKAENEGKFSFAPV 459
DB 337 DNIDITGMLGYPMQIKVFLCRDSTLAPVLDLALFLDLAKLEMRGIGQM----- 388
QY 460 ATISLYLTKAPLVPPGTPVNVALSQRAMLENIRACVGLAPENNMILEY 509
DB 369 ---MSFYKFTMAQGLPVEHDLFIQLTKMLNTLKVAGGEPTITLGLDY 435

RESULT 8
US-09-107-433-4707
; Sequence 4707, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: Nucleic Acid and Amino Acid
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
```


NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4707:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURES:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...730
SEQUENCE DESCRIPTION: SEQ ID NO: 4707:
US-09-107-433-4707
Query Match 4.3%; Score 113.5; DB 4; Length 730;
Best Local Similarity 20.8%; Pred. No. 0.058;
Matches 108; Conservative 64; Mismatches 183; Indels 165; Gaps 29;
QY 90 WATKDKIQOANYFG-----SLTQASAIRVGSFGGEIYAPFKSLPMPNPPDIYFGMD 143
DB 218 WGEVDASKR--YFGVSASEVSLDQAATL-AGMLKGELYNPLNSVEDSTNRDVT----- 269
QY 144 ISNMNLA--DAMARAKVFDIDLQKQLRPYMESMLP---LPGIYDPDFIAA-----NOE 191
DB 270 LQNMVAAGYIDKNGQETEALEVDMTSQLDHDKYEGKISDYRYPSPFPAVVAEAVSKVLTBE 329
QY 192 EKAN--VIKGTQEQVQOIIKDIKAFKA-----TKYDKVYVLM 229
DB 330 EIVNNGYRIYTELDQNYQANMOIYVENISLFPRAEDGTPAQSVALLEPKTGVGRGVQ 389
QY 230 TANTER-----YSN-----LVV-----GNDTMENTLAADVREAEI 261
DB 390 VADNDKGTGRNRYVATQSKRSPGKTIKPLVVTTPAVEAGMALNKQDNTHTMQYDSYKV- 448
QY 262 SPSTLYALACVMEVNPFGSPONTFVPGLI---DLAIARNTLIGDDPKSGQTKKSVL 318
DB 449 ---NYAGIKTSREVMYQALAESLNLPAVATVNDLGVDK-APEAGEKFGLANMEKVDRL 503
QY 319 VDFLVGAGIK--PTSIIVSNHGNNDG-----NLGAPQTFRSKEISKSNVVD 365
DB 504 -GVALGSGVETNPLQMAQVAAFAFANEGMLPEAHFISRIENASG-QVIANHKNSQKRVIDK 561
QY 366 WV-----NSNALIYEPGEHPDHVVVIVKVPYVGSKRAMD-EYTSSEIFMG- 410
DB 562 SVADKMTSMMLGTFTNGTGISSSPA---DYVMAK---TGTTAEVFNDEYISDQWVICY 614
QY 411 -KSTIVLH-----NFCEDSLAAPITLDVLVLAELSTRIEFAENEGKFSHPVA-TTL 463
DB 615 TDDVVISHLGFPPTDENHYLAG-----STSGAAHVFRNIAITIL 655
QY 464 SYLTKAPLVPGTPVYNALSKORAMLENIMRACVGLAPEN 503
DB 656 -----PYTPGSTFTVENAYKON-----GIAPAN 678
RESULT 9
US-09-583-110-4720
Sequence 4720, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 4720
LENGTH: 731
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-4720
Query Match 4.3%; Score 113.5; DB 4; Length 731;
Best Local Similarity 20.8%; Pred. No. 0.058;
Matches 108; Conservative 64; Mismatches 183; Indels 165; Gaps 29;
QY 90 WATKDKIQOANYFG-----SLTQASAIRVGSFGGEIYAPFKSLPMPNPPDIYFGMD 143
DB 215 WGEVDASKR--YFGVSASEVSLDQAATL-AGMLKGELYNPLNSVEDSTNRDVT----- 266
QY 144 ISNMNLA--DAMARAKVFDIDLQKQLRPYMESMLP---LPGIYDPDFIAA-----NOE 191
DB 267 LQNMVAAGYIDKNGQETEALEVDMTSQLDHDKYEGKISDYRYPSPFPAVVAEAVSKVLTBE 326
QY 192 EKAN--VIKGTQEQVQOIIKDIKAFKA-----TKYDKVYVLM 229
DB 327 EIVNNGYRIYTELDQNYQANMOIYVENISLFPRAEDGTPAQSVALLEPKTGVGRGVQ 386
QY 230 TANTER-----YSN-----LVV-----GNDTMENTLAADVREAEI 261
DB 387 VADNDKGTGRNRYVATQSKRSPGKTIKPLVVTTPAVEAGMALNKQDNTHTMQYDSYKV- 445
QY 262 SPSTLYALACVMEVNPFGSPONTFVPGLI---DLAIARNTLIGDDPKSGQTKKSVL 318
DB 446 ---NYAGIKTSREVMYQALAESLNLPAVATVNDLGVDK-APEAGEKFGLANMEKVDRL 500
QY 319 VDFLVGAGIK--PTSIIVSNHGNNDG-----NLGAPQTFRSKEISKSNVVD 365
DB 501 -GVALGSGVETNPLQMAQVAAFAFANEGMLPEAHFISRIENASG-QVIANHKNSQKRVIDK 558
QY 366 WV-----NSNALIYEPGEHPDHVVVIVKVPYVGSKRAMD-EYTSSEIFMG- 410
DB 559 SVADKMTSMMLGTFTNGTGISSSPA---DYVMAK---TGTTAEVFNDEYISDQWVICY 611
QY 411 -KSTIVLH-----NFCEDSLAAPITLDVLVLAELSTRIEFAENEGKFSHPVA-TTL 463
DB 612 TDDVVISHLGFPPTDENHYLAG-----STSGAAHVFRNIAITIL 652
QY 464 SYLTKAPLVPGTPVYNALSKORAMLENIMRACVGLAPEN 503
DB 653 -----PYTPGSTFTVENAYKON-----GIAPAN 675
RESULT 10
US-09-489-039A-11570
Sequence 11570, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Bretton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 11570
LENGTH: 842
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11570

Query Match 4.3%; Score 112.5; DB 4; Length 842;
Best Local Similarity 22.9%; Pred. No. 0.092;
Matches 100; Conservative 63; Mismatches 139; Indels 135; Gaps 27;

QY 30 ELVHENRNGTYWIKPKSVKTEFKNTIHVPEKLGVMYWGCGNNSTLTGCVIANREGIS 89
DB 426 QLVDDOVQDATAAYL---KGIDPD-----NPPSGDGFINAFYREGT----- 463
QY 90 WATKXKIOQANVFGLTQASAIRVGSFQGEELIYAPFKSLPMVNPDDIVFGMD:SNM-- 147
DB 464 -----QRFNDFENLRVAGQY-IGSFPR---TMTPTASTQQYVS-----DLKNQV 502
QY 148 NLADAMARAKYVD-----IDLQKOLRPYMESMLPLPGIY-----DPDEFIANOEER 193
DB 503 NQEGLAGAARAAFDHQAARLISQKQSDPF-QSAVDI-GAYKPISTNPDAIABEVNR 560
QY 194 --ANNVIRG-----TKQEOVOQI IKDKAFKKAATKDKVAVLMTANTERYSNLVVGL 243
DB 561 YAAQDQLKAIIGITPPLLSKQF--SQVLTQ--AVRSTVDVNCAL-----SLQGL 605
QY 244 NDTM-ENLLAAYDRNEAISPSTLYALCV-MENVPTINGS---PONTFV---PGLIDLA 295
DB 606 GRTLPKALRSVASSIAPSGPCTAYALALLGQDDQYDRSGIITPSQFVSFKPTLDDKD 665
QY 236 IARNTLIGDDPFKSGQTKKSVLVDFLVGAGIKPTSIYSYNHL-----GNNDGMNLS 347
DB 666 VAK-TVLAGDQMLNPTKMKD-----AGISAVSIPDEKLTQYFDKEVGNAPAYSQ 716
QY 348 APO-----TFRSKEISKSNNV--DDMNNSNALIYEPGE-----HPDHVVYIK 387
DB 717 ARQVAMGNFRSAYALAYOSGDSKNTIVSPSDIAEKAAQYATGAVYKGLGSDV--- 773
QY 388 YVPYVGDSCRAMDEXYS 404
DB 774 -MPFGMDKTTFRDRYTA 789

RESULT 11
US-09-248-796A-20645
Sequence 20645, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 20645
LENGTH: 452
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-20645

Query Match 4.2%; Score 111.5; DB 4; Length 452;
Best Local Similarity 20.6%; Pred. No. 0.041;
Matches 105; Conservative 83; Mismatches 157; Indels 165; Gaps 31;

QY 11 EVVXKTER---EIGSYVNYE-TTELVH---ENRNGTYWIKPKSVKTEFK----- 54
DB 2 PSIVVPDSTTKOYQAMNSENATEEYPILESK-----IVMLDENYILKIILOSIVAN 55
QY 55 -TNIHVPKLGVMLV---GMGNGNGSTLTGVIANREGISWATKXKIOQANV-FGSLTQA 108
DB 56 HPTITINQIPMLLIPSLTWSROSIEYITKYIENLEITAFNIIIDLSLAATGVCQLTNS 115
QY 109 SAIRVGSFQGEELIYAPFKSLPMVNPDDIVFGMD:SNM-----MNL-----A 150

DB 116 TVVYV---DDENI-----QIVPVVGYQAIKFKAGLKIKNGSITTSREIKQNLPNLTSCQI 167
QY 151 DAMARAKYVD--IDLQKOLRPYMESMLPLPGIYDPDF-----IANOEERANNVIKGTQ 203
DB 168 EDLKNSDIFEEVVIDQOGVLDYIKDITKTNNDEDFNEFVAKIVENQNGIEAIIISNPT 227
QY 204 EOVOQI IKDI-KAFKKAATK---VDKVV--VLMTANTERYSNLVVGLNDTMENLLAAVDNR 257
DB 228 EQQKQEQGDSKPKVNELEKKNYFIDSKTQEKIWKIG-ERFS---GTN-----N 270
QY 258 EAELSPSTLYAIACMENVPTINGSPONTFVPGGLIDLAIAANTLIGDDPFKSGQTKMSV 317
DB 271 LVKLISSEIYS---SLSTIPID-KRQDCY-----DNIIIVGSIFFTPGLK-BAI 315
QY 318 LV---DLYV-----GAGI-----KPTSIYSYNHLGNNDGMNLSAAPTFRS 354
DB 316 LIKANQDYLVAEPNALIDOSINDPGVNTAILKYQOSTITINDINEGSGDNN----- 366
QY 355 KEISKSNVDDMVNSNALIYEPGEHPDHVVVIXYVPYVGDSCRAMDE----- 401
DB 367 ---SNSN-----SNSNQV-----PNSIKLVKYPDYPEPKKPKREKGSWHDYFLGQ 411
QY 402 -YTSRIFMG-----GKSTIVLHNTCEDSL 425
DB 412 IYSKQIYSGSHHGKELFVGSWMYERVL 441

RESULT 12
US-09-014-897-2
Sequence 2, Application US/09014897
Patent No. H002085
GENERAL INFORMATION:
APPLICANT: Hoskins, JoAnn
Jaskunas, S. Richard
Rockey, Pamela K.
Zhao, Genshi
Rostock, Paul R. Jr.
No. H002085rfs, Franklin H.
TITLE OF INVENTION: Penicillin Binding Protein From
Streptococcus Pneumoniae
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,897
FILING DATE: 28-Jan-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/731,716
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-10,887
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 731 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-014-897-2

Query Match 4.2%; Score 111.5; DB 1; Length 731;
Best Local Similarity 20.8%; Pred. No. 0.091;
Matches 108; Conservative 63; Mismatches 184; Indels 165; Gaps 29;

QY 90 WATKXIOQANYFG-----SLTQSAIRVGSFQGEETIAPFKSLPMPNPPDIYFGGWD 143
DB 215 WGEVDSKX--YFGVSASBSVSLDQAATL--AGMLKGPGLYNPLNSVEDSTNRDVTV----- 266
QY 144 ISNMMLA---DAMARAKVFDIDLQKQRLPYMESMLP---LPGIYDDPFIYA-----NOE 191
DB 267 LQNMVAAGIDKNGQTEAEAVDMTSQLHDKYSGKISDRYRPSYFDVAVNEAVSKNLTBE 326
QY 192 ERANN-----VIKTKQEQVOQIIKDIAKFEA-----TYDKVYVILM 229
DB 327 EIVNNGYRIYETLDQNYQANMOIYVENTSLFPRAEDGTFAGSGVALBPKTGVRGVGQ 386
QY 230 TANTER-----YEN-----LVV-----GINDMENLLAVDRNEAI 261
DB 387 VADNDKTRFNRYATQSKRSPGSTIKPLVYTPPAVEAGMALNKOLDHNTMQDYKYVD- 445
QY 262 SPSTLYAIAQWENVPFINGSPONTFVPGLI---DLAIARNTLIGDDPKSGQTKMKSVL 318
DB 446 ---NYAGIKTSREVPMTQSLAESLNLPAVATVNDLGYDK-AFEAGEKFGLMMEKYDRL 500
QY 319 VDFLVAGIK--PTSIYSVYHNLGNNDGM-----NLSAPQTFRSKEISKSNNVDD 365
DB 501 -GVALGSGVETNPLQMAQYAAFANEGMLPEAFISRIENASG-QVIAHKNSQKRVIDK 558
QY 366 NV-----NSNALYEPGEHPDHVVIVIKYVPVGDGSKAMD-EYTSSEIFMG- 410
DB 559 SVADKMTSMMLGTFTNGTGISSPA---DYWAGK---TGTTAEVFNPEYTSDDQVIGY 611
QY 411 -KSTIVLH-----NTCEDSLAAPITLDVLLAELSTRIEFKAENEGKFSHPVA-TTL 463
DB 612 TPDVVISHMLGFPPTDENHYLAG-----STNGAAHVFRNIANTIL 652
QY 464 SYLTAPLVPGPTPVVNALSKORAMLENIMRACVGLAPEN 503
DB 653 -----PYTPGSTFTVENAYKQN-----GIAPAN 675

RESULT 13
US-08-731-716-2
; Sequence 2, Application US/08731716
; Patent No. 5789202
; GENERAL INFORMATION:
; APPLICANT: Hoskins, Johann
; APPLICANT: Jaskunas, S. Richard
; APPLICANT: Rokey, Pamela K.
; APPLICANT: Zhao, Genehl
; APPLICANT: Rostock, Paul R. Jr.
; APPLICANT: No. 5789202ris, Franklin H.
; TITLE OF INVENTION: Penicillin Binding Protein From
; TITLE OF INVENTION: Streptococcus Pneumoniae
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,716
; FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-10,887
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 731 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-731-716-2

Query Match 4.2%; Score 111.5; DB 1; Length 731;
Best Local Similarity 20.8%; Pred. No. 0.091;
Matches 108; Conservative 63; Mismatches 184; Indels 165; Gaps 29;

QY 90 WATKXIOQANYFG-----SLTQSAIRVGSFQGEETIAPFKSLPMPNPPDIYFGGWD 143
DB 215 WGEVDSKX--YFGVSASBSVSLDQAATL--AGMLKGPGLYNPLNSVEDSTNRDVTV----- 266
QY 144 ISNMMLA---DAMARAKVFDIDLQKQRLPYMESMLP---LPGIYDDPFIYA-----NOE 191
DB 267 LQNMVAAGIDKNGQTEAEAVDMTSQLHDKYSGKISDRYRPSYFDVAVNEAVSKNLTBE 326
QY 192 ERANN-----VIKTKQEQVOQIIKDIAKFEA-----TYDKVYVILM 229
DB 327 EIVNNGYRIYETLDQNYQANMOIYVENTSLFPRAEDGTFAGSGVALBPKTGVRGVGQ 386
QY 230 TANTER-----YEN-----LVV-----GINDMENLLAVDRNEAI 261
DB 387 VADNDKTRFNRYATQSKRSPGSTIKPLVYTPPAVEAGMALNKOLDHNTMQDYKYVD- 445
QY 262 SPSTLYAIAQWENVPFINGSPONTFVPGLI---DLAIARNTLIGDDPKSGQTKMKSVL 318
DB 446 ---NYAGIKTSREVPMTQSLAESLNLPAVATVNDLGYDK-AFEAGEKFGLMMEKYDRL 500
QY 319 VDFLVAGIK--PTSIYSVYHNLGNNDGM-----NLSAPQTFRSKEISKSNNVDD 365
DB 501 -GVALGSGVETNPLQMAQYAAFANEGMLPEAFISRIENASG-QVIAHKNSQKRVIDK 558
QY 366 NV-----NSNALYEPGEHPDHVVIVIKYVPVGDGSKAMD-EYTSSEIFMG- 410
DB 559 SVADKMTSMMLGTFTNGTGISSPA---DYWAGK---TGTTAEVFNPEYTSDDQVIGY 611
QY 411 -KSTIVLH-----NTCEDSLAAPITLDVLLAELSTRIEFKAENEGKFSHPVA-TTL 463
DB 612 TPDVVISHMLGFPPTDENHYLAG-----STNGAAHVFRNIANTIL 652
QY 464 SYLTAPLVPGPTPVVNALSKORAMLENIMRACVGLAPEN 503
DB 653 -----PYTPGSTFTVENAYKQN-----GIAPAN 675

RESULT 14
US-09-512-250C-33
; Sequence 33, Application US/09512250C
; Patent No. 6518042
; GENERAL INFORMATION:
; APPLICANT: Borchert, Torben
; APPLICANT: Pedersen (Executor for Lars Christiansen, deceased), Dennis
; APPLICANT: Vind, Jesper
; TITLE OF INVENTION: A process for making DNA libraries in filamentous fungal cells
; TITLE OF INVENTION: Cells
; FILE REFERENCE: 5718.200-US
; CURRENT APPLICATION NUMBER: US/09/512,250C
; CURRENT FILING DATE: 1999-02-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 33

LENGTH: 935
TYPE: PRT
ORGANISM: mus. p.
US-09-512-250C-33

Query Match 4.2%; Score 111; DB 4; Length 935;
Best Local Similarity 18.9%; Pred. No. 0.15;
Matches 100; Conservative 76; Mismatches 218; Indels 134; Gaps 19;

```
QY 52 EKKTHHPKLVGMVGMGN-----GSLTGG-----VIANREGISMARD 94
DB 177 EPEPDQSNLEALLIQIGPKCVLPGGETTDMGKLRQVIORGILLTERRADFTKD 236
QY 95 KIQQANYFGSLTQASAIRVGSFGEEIYAPFKSLPMVN-----PDDIVFGMDISN--- 146
DB 237 IYQDLNRLKGGKQGINSAALPEMENQVAVSSLSAVIKFLELLSDDSFGFELATPDF 296
QY 147 ---NNIADAMARA-----KVFIDLOKQLRPEYMESMLPLPG-----IYDPFIANAQ 190
DB 297 SQYMLDMAAVRALNLPQGSVEDTGSGLAALLNKCKTAQOQRLVNGWIKQPLMDRRI 356
QY 191 EERANNVIKGTQEOVOQIIRK--IKAPKEATKVDKVVVLTANTERYSNLVVGNDTME 248
DB 357 EERLNLVEAFVEDSELROSLQEDLLRRFPDLNRLAKKQORQANLQDCYRLQGINQ--LP 415
QY 249 NLLAAVDRNEAISPSTLYAIACV-----MENVPFINGSPQ 284
DB 416 SVIQALEKEGR--HQALLAVFVTPILDRSPSKQEMIEETLMDQVENHEFL---VK 471
QY 285 NTFVPGILDIALARTLLIGDDFKSGQTKKSVLVDFLVGAGIKPTSIYSYHNLGNDGM 344
DB 472 PEFDPNLSELREVMGDL-----EKKQSTLINAAGLGLDP-----GKQIKL 513
QY 345 NLSAP-----QTFPSKEISKN-----VYDDMVNSNALIYERGEHPDHVVVIKVPYVG 394
DB 514 DSSAQFGYFRVYTCCKEKKLRNNKNFSTVD--IQKGVFTNSELS--LNEEYTKNKE 569
QY 395 SKRAMDEYTSEIFMGKSTIVLHNTCEDSLAPILDLVLAELESTRIEFKRANEKGPH 454
DB 570 YEAQDAIYKEIVNISSGVEMQTLND-----VLAHDAIYVF----- 608
QY 455 SFHPVATIIISYLTKAPLYPFGTPPVNALSQRAMLENTIRACVGLAPE 502
DB 609 -----AHVSNAPVYVPRVILEKGRRIILKASHRACVEVODE 647
```

RESULT 15
US-09-583-110-3066
Sequence 3066, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
FILE REFERENCE: PATH00-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 3066
LENGTH: 747
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-3066

Query Match 4.2%; Score 109.5; DB 4; Length 747;
Best Local Similarity 21.3%; Pred. No. 0.15;
Matches 102; Conservative 56; Mismatches 161; Indels 159; Gaps 25;

```
QY 2 FIENFKVECPNVKYTEIEIQSVNYVETTELHENRNGTYQIIVKRSKYEFKTNIAVHK 61
DB 173 FIENFGVHTTVSDIQNSIVTFYDVHTDIFRLN-----KLNIDISE 214
QY 62 LGVNLVGMGNNGSLTGGVIANREGISMATKDKIQQANYFGSLTQASAIRVGSFGSEI 121
DB 215 ANIM-----NQTT---VILDEKNIELL---LSKAPLYVSM-----I 244
QY 122 YAPFSLIPMVNPPDIVFGMDISNMNLADAMARAKVFDILOKQLRPEYMESMLPLPGIY 181
DB 245 VEDFSKL---SVDDFSLDNDL--KINLPSMNEPVGVIDTLDPKRYFNEBWYEHDFV 299
QY 182 DPDFFIANQE-----ERAN---NYIKGTQEOVOQI-----IKDIAF 216
DB 300 SPDISKSDQDYKHGTAVTSLVLDGANLNPNLDDGCGNFRVHFGVSLQSGFNSFTIIKQI 359
QY 217 KEATKVDKVVVLTANTERYSNLVVGNDTM--ENLLAAVDR--NEAETSPSTLYAIACVM 273
DB 360 KEIVSQNADIKVM-----NLSIGSNDEIRENFISAEGALLDEIOPENDVIFIIAGT- 410
QY 274 ENVPFINGSPQNTFVPGILDIALARTLLI--GGDF--KSGQTKKSVLVDFLVGAGIKPT 330
DB 411 -NASVINKRKRKIGAP---ADSLNSTIIVNSVDPNNQSVISREGIVLSFFV---KPD 460
QY 331 SIVSYNHLGND-----GNNLSAPQTFR-----SKEISKNVVD- 364
DB 461 --VSYYGGNDDFINVCPRLGIRVAGTSFAAPFIARGAVLIHIMGLSREBAKALLIDA 518
QY 365 -----DM-----VNSNALIYERGEHPDHVVVIKVPYVGDSKRAMDEY 402
DB 519 AIPWMDKKTFTDLSLIGNGIIVPIKMDILSTPDDE-----IKFI--VSDISRAYDTY 568
```

Search completed: June 7, 2005, 16:48:53
Job time : 35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 03:03:45 ; Search time 1094.05 Seconds

(without alignments)
9972.402 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760

Sequence: 1 cctctcttattccttctgt.....aatgtagtataatttctgt 1760

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*

15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*

20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*

21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*

22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:*

23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*

24: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*

25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1760	100.0	1760	US-10-025-003-1	Sequence 1, Appli
2	1760	100.0	1760	US-10-718-952-1	Sequence 1, Appli
3	1745.8	99.2	1989	US-10-424-599-70167	Sequence 1, Appli
4	1533	87.1	1533	US-10-025-003-9	Sequence 9, Appli
5	1533	87.1	1533	US-10-718-952-9	Sequence 9, Appli
6	1531.4	87.0	1533	US-10-025-003-5	Sequence 5, Appli
7	1531.4	87.0	1533	US-10-718-952-5	Sequence 5, Appli

8	1465.8	83.3	1533	14	US-10-025-003-13	Sequence 13, Appl
9	1465.8	83.3	1533	14	US-10-025-003-15	Sequence 15, Appl
10	1465.8	83.3	1533	19	US-10-718-952-13	Sequence 13, Appl
11	1465.8	83.3	1533	19	US-10-718-952-15	Sequence 15, Appl
12	1464.2	83.2	1533	14	US-10-025-003-11	Sequence 11, Appl
13	1464.2	83.2	1533	19	US-10-718-952-11	Sequence 11, Appl
14	1296.4	73.7	2018	18	US-10-424-599-12022	Sequence 12022, A
15	1136.8	64.6	2582	18	US-10-424-599-12021	Sequence 12021, A
16	1007	57.2	1824	19	US-10-767-701-14714	Sequence 14714, A
17	1000.6	56.9	1941	20	US-10-739-930-2091	Sequence 2091, Ap
18	999.2	56.8	2576	19	US-10-437-963-70769	Sequence 70769, Ap
19	995.8	56.6	2284	20	US-10-425-115-47189	Sequence 47189, A
20	995.4	56.6	1533	9	US-09-938-842A-1438	Sequence 1438, Ap
21	995.4	56.6	1533	11	US-09-938-842A-1438	Sequence 1438, Ap
22	994.6	56.5	2582	18	US-10-424-599-12021	Sequence 12021, A
23	994.2	56.5	1879	18	US-10-425-114-5704	Sequence 5704, Ap
24	994.2	56.5	1912	18	US-10-425-114-5704	Sequence 5704, Ap
25	994.2	56.5	1914	18	US-10-425-114-5612	Sequence 5612, Ap
26	994.2	56.5	1959	9	US-09-727-628-1	Sequence 1, Appli
27	994.2	56.5	1997	20	US-10-425-115-47190	Sequence 47190, A
28	994.2	56.5	2156	18	US-10-425-114-5858	Sequence 25858, A
29	993.8	56.5	2309	20	US-10-425-115-47194	Sequence 47194, A
30	992.6	56.4	1971	18	US-10-425-114-26983	Sequence 26983, A
31	992.6	56.4	2233	20	US-10-425-115-47187	Sequence 47187, A
32	991	56.3	2444	20	US-10-425-115-122151	Sequence 122151, A
33	990.4	56.3	1898	18	US-10-425-114-7081	Sequence 7081, Ap
34	984	55.9	1931	9	US-09-921-232-10	Sequence 10, Appl
35	984	55.9	1931	9	US-09-921-330-10	Sequence 10, Appl
36	984	55.9	1931	9	US-09-921-329-10	Sequence 205, App
37	956.4	54.3	1759	19	US-10-363-8529-205	Sequence 47180, A
38	920.8	52.3	1673	20	US-10-425-115-47180	Sequence 14009, A
39	919	52.2	1858	18	US-10-425-114-14009	Sequence 5864, Ap
40	919	52.2	2050	18	US-10-425-114-1864	Sequence 18524, A
41	917.4	52.1	1853	18	US-10-425-114-16524	Sequence 83553, A
42	917.4	52.1	2130	20	US-10-425-115-83553	Sequence 47182, A
43	906.4	51.5	1789	20	US-10-425-115-47182	Sequence 85785, A
44	904.6	51.4	2196	19	US-10-437-963-85285	Sequence 5619, Ap
45	676.6	38.4	1240	18	US-10-425-114-5619	

ALIGNMENTS

RESULT 1

US-10-025-003-1

; Sequence 1, Application US/10025003

; Publication No. US2003007465A1

GENERAL INFORMATION:

APPLICANT: Hitz, William

APPLICANT: Sebastian, Scott

APPLICANT: Grace, John

APPLICANT: Strelt, Leon

TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE

FILE REFERENCE: BB-1077-C

CURRENT APPLICATION NUMBER: US/10/025,003

CURRENT FILING DATE: 2002-05-07

PRIOR APPLICATION NUMBER: 08/835,751

PRIOR FILING DATE: APRIL 8, 1997

PRIOR APPLICATION NUMBER: PCT/US98/06822

NUMBER OF SEQ ID NOS: 16

PRIOR FILING DATE: APRIL 7, 1998

SOFTWARE: Microsoft Office 97

SEQ ID NO 1

LENGTH: 1760

TYPE: DNA

ORGANISM: Glycine max

US-10-025-003-1

Query Match 100.0%; Score 1760; DB 14; Length 1760;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCTTTATTCCTTTGTAAATTCATTCATCTTAACTTTGTGAAAAATATGTCA 60
| | | | |
Db 1 CTTCTTTATTCCTTTGTAAATTCATTCATCTTAACTTTGTGAAAAATATGTCA 60
QY 61 TCGAAGATTTTAAAGTTGAGTCTCTAATGTGATACCGAGACTGAGATTCAGTCCG 120
| | | | |
Db 61 TCGAAGATTTTAAAGTTGAGTCTCTAATGTGATACCGAGACTGAGATTCAGTCCG 120
QY 121 TGTACAACTAGAGAAACCAACGAGCTTTGTCAGAGAACGAGAAATGGACCTACAGTGA 180
| | | | |
Db 121 TGTACAACTAGAGAAACCAACGAGCTTTGTCAGAGAACGAGAAATGGACCTACAGTGA 180
QY 181 TTGTCAAAACCCAAATCTGTCAAAATAAGAAATTTAAACCAACATCCATGTTCTTAAATAG 240
| | | | |
Db 181 TTGTCAAAACCCAAATCTGTCAAAATAGAAATTTAAACCAACATCCATGTTCTTAAATAG 240
QY 241 GGGTAAATGTTGGGTGGGTGGGTGGAAACACAGGCTCAACCTCAACGGTGGTGTATTTG 300
| | | | |
Db 241 GGGTAAATGTTGGGTGGGTGGGTGGAAACACAGGCTCAACCTCAACGGTGGTGTATTTG 300
QY 301 CTAAACGAGAGGGGCAATTCATGGGCTACAAAGAGCAAGATTCAAACAGCCAAATTAATTG 360
| | | | |
Db 301 CTAAACGAGAGGGGCAATTCATGGGCTACAAAGAGCAAGATTCAAACAGCCAAATTAATTG 360
QY 361 GCTCCCTCAACCAAGCCTCAGCTATCCAGATTGGGTCCTTCCAGGAGAGAGAAATCTATG 420
| | | | |
Db 361 GCTCCCTCAACCAAGCCTCAGCTATCCAGATTGGGTCCTTCCAGGAGAGAGAAATCTATG 420
QY 421 CCCCATTCAAAGCCTGCTTCCAAATGTGTTAACCTGACACATTTGTGTTGGGGGATGGG 480
| | | | |
Db 421 CCCCATTCAAAGCCTGCTTCCAAATGTGTTAACCTGACACATTTGTGTTGGGGGATGGG 480
QY 481 AATATAGCAACATGAACCTGGCTGATGACATGAGCCGAGGCAAGGTTGTGACATGAT 540
| | | | |
Db 481 AATATAGCAACATGAACCTGGCTGATGACATGAGCCGAGGCAAGGTTGTGACATGAT 540
QY 541 TGCAGAGAGAGTTGAGGCTTATCATGAAATCCATGCTTCCATCCCGGAAATCTATGACC 600
| | | | |
Db 541 TGCAGAGAGAGTTGAGGCTTATCATGAAATCCATGCTTCCATCCCGGAAATCTATGACC 600
QY 601 CGGATTTTCATTCGCTGCCAACCAAGAGAGCGTCCCAACAGTCAATCAAGGCAACAAAGC 660
| | | | |
Db 601 CGGATTTTCATTCGCTGCCAACCAAGAGAGCGTCCCAACAGTCAATCAAGGCAACAAAGC 660
QY 661 AAGACAAGTTCAACAAATCATCAAGAGATCATGAGCGCTTTAAAGGAAGCCACCAAGTGG 720
| | | | |
Db 661 AAGACAAGTTCAACAAATCATCAAGAGATCATGAGCGCTTTAAAGGAAGCCACCAAGTGG 720
QY 721 ACAAGGTGGTTGACTGTGAGCTGCTCCACACAGAGAGGTACAGTAATTTGTGTGGGCC 780
| | | | |
Db 721 ACAAGGTGGTTGACTGTGAGCTGCTCCACACAGAGAGGTACAGTAATTTGTGTGGGCC 780
QY 781 TTAATGACACCATGAGAAATCTTTGGCTGCTGTGAGACGAATATGAGGTGATTTCTC 840
| | | | |
Db 781 TTAATGACACCATGAGAAATCTTTGGCTGCTGTGAGACGAATATGAGGTGATTTCTC 840
QY 841 CTTCCACCTTGTATGCAATTCCTTGTGTTAAGAAAAATTTCTTTCATTAATGAGAGCC 900
| | | | |
Db 841 CTTCCACCTTGTATGCAATTCCTTGTGTTAAGAAAAATTTCTTTCATTAATGAGAGCC 900
QY 901 CTCAAGAACATTTTGTACCAAGGCTGATTAATCTTCCATCCGAGAGAACACTTTGATTTG 960
| | | | |
Db 901 CTCAAGAACATTTTGTACCAAGGCTGATTAATCTTCCATCCGAGAGAACACTTTGATTTG 960
QY 961 GTGAGATGATCTTCAAGATGCTGACAGCAAAATGAAATCTGTGTTGTTGATTTCTTTG 1020
| | | | |
Db 961 GTGAGATGATCTTCAAGATGCTGACAGCAAAATGAAATCTGTGTTGTTGATTTCTTTG 1020
QY 1021 TGGGGGCTGGTATCAAGCAACATCATATGATTAAGATTACCAACATCTGGGAAAAATATGATG 1080
| | | | |
Db 1021 TGGGGGCTGGTATCAAGCAACATCATATGATTAAGATTACCAACATCTGGGAAAAATATGATG 1080
QY 1081 GTATGAATCTTTCGGCTCCACAACTTTCGTTCCAGAGAAATCTCCAGAGCAACGTTG 1140
| | | | |

Db 1081 GTATGAATCTTTCGGCTCCACAACTTTCGTTCCAGAGAAATCTCCAGAGCAACGTTG 1140
QY 1141 TTGATGATATGATCAACAGCAATGCCATCTCTATGAGCCTGGTGAATCATCAGACATG 1200
| | | | |
Db 1141 TTGATGATATGATCAACAGCAATGCCATCTCTATGAGCCTGGTGAATCATCAGACATG 1200
QY 1201 TTGTTGTTAATTAAGATATGATGCTTACGTAGGAGGACAGCAAGAGCCATGATGATGACA 1260
| | | | |
Db 1201 TTGTTGTTAATTAAGATATGATGCTTACGTAGGAGGACAGCAAGAGCCATGATGATGACA 1260
QY 1261 CTTACAGATATTTCAATGGGTGGAAAGAGACACATTTGTTGACAAACATGCGAGATT 1320
| | | | |
Db 1261 CTTACAGATATTTCAATGGGTGGAAAGAGACACATTTGTTGACAAACATGCGAGATT 1320
QY 1321 CCGCTTATGCTGCTCTATTAATCTTGAACCTTGATCTTCTTCTGAGCTCAGACTAGAA 1380
| | | | |
Db 1321 CCGCTTATGCTGCTCTATTAATCTTGAACCTTGATCTTCTTCTGAGCTCAGACTAGAA 1380
QY 1381 TCGAGTTTAAAGCTGAAATAGAGGAAATTCACACTATTCACCCAGTTGCTACATCC 1440
| | | | |
Db 1381 TCGAGTTTAAAGCTGAAATAGAGGAAATTCACACTATTCACCCAGTTGCTACATCC 1440
QY 1441 TCAGTACCTCAACCAAGGCTCTCTGCTTCCACCGGGTACACAGTGTGAATGATTTGT 1500
| | | | |
Db 1441 TCAGTACCTCAACCAAGGCTCTCTGCTTCCACCGGGTACACAGTGTGAATGATTTGT 1500
QY 1501 CAAAGCAGCGTGAATGCTGAGAAAACATATAGAGGCTGTGTTGGAATTTGAGCCACAGAGA 1560
| | | | |
Db 1501 CAAAGCAGCGTGAATGCTGAGAAAACATATAGAGGCTGTGTTGGAATTTGAGCCACAGAGA 1560
QY 1561 ATTAACATGATCTCGAGTACCAAGTGAAGCATGAGACCGAAGAAATATATAGTTGGGGTAG 1620
| | | | |
Db 1561 ATTAACATGATCTCGAGTACCAAGTGAAGCATGAGACCGAAGAAATATATAGTTGGGGTAG 1620
QY 1621 CCTAGCTGAATGTTTATGTTAATATATATGTTGCTTAAATTTTGGCAAGGTATTTGAA 1680
| | | | |
Db 1621 CCTAGCTGAATGTTTATGTTAATATATATGTTGCTTAAATTTTGGCAAGGTATTTGAA 1680
QY 1681 TGCATCAGCTTCATTAATGCTTTAGAGCGGGGCAATATCTGTTTACTAGAAACATGAATG 1740
| | | | |
Db 1681 TGCATCAGCTTCATTAATGCTTTAGAGCGGGGCAATATCTGTTTACTAGAAACATGAATG 1740
QY 1741 AATGATATATATTTTGTGT 1760
| | | | |
Db 1741 AATGATATATATTTTGTGT 1760
| | | | |
RESULT 2
US-10-718-952-1
; Sequence 1, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACCCHARIDES AND PHYTIC ACID
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1760
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-1

Query Match 100.0%; Score 1760; DB 19; Length 1760;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CTCTTTTATTCCTTTTGTATTTTCAATTCATTCTTAATCTTTGTGAAAAATATATGTC 60
DB 1 CTCTTTTATTCCTTTTGTATTTTCAATTCATTCTTAATCTTTGTGAAAAATATATGTC 60
QY 61 TCGAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGATTCACTCCG 120
DB 61 TCGAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGATTCACTCCG 120
QY 121 TGTACAACTACGAAACCAACCGAATCTTTCACGAGAAACGAAATGCACTATCAGTGA 180
DB 121 TGTACAACTACGAAACCAACCGAATCTTTCACGAGAAACGAAATGCACTATCAGTGA 180
QY 181 TTGTCAAAACCAATGTCTCAAAATAGCAATTTAAAAACAATCCATCTGTTCTTAATAG 240
DB 181 TTGTCAAAACCAATGTCTCAAAATAGCAATTTAAAAACAATCCATCTGTTCTTAATAG 240
QY 241 GGGTAAATGCTTGTGGGTGGGTGGAACAACGGCTCAACCTCACCGGTGTGTTATTG 300
DB 241 GGGTAAATGCTTGTGGGTGGGTGGAACAACGGCTCAACCTCACCGGTGTGTTATTG 300
QY 301 CTAAACGAGAGGCAATTTTCAATGGGCTTACAAAGACAAAGATTTCAACAGCCAAATTA 360
DB 301 CTAAACGAGAGGCAATTTTCAATGGGCTTACAAAGACAAAGATTTCAACAGCCAAATTA 360
QY 361 GCTCCCTCAACCAAGCCTCAGCTATCCGAGTTGGGTCTTTCAGAGGAGAAATCTATG 420
DB 361 GCTCCCTCAACCAAGCCTCAGCTATCCGAGTTGGGTCTTTCAGAGGAGAAATCTATG 420
QY 421 CCCCATTTCAAGAGCCTGCTTCAATGTTAACTTGAACAATTTGTTGGGGAATGG 480
DB 421 CCCCATTTCAAGAGCCTGCTTCAATGTTAACTTGAACAATTTGTTGGGGAATGG 480
QY 481 ATATCGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTTTGAATGATG 540
DB 481 ATATCGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTTTGAATGATG 540
QY 541 TCGAGAGAGGTTGAGGCTTATCAATGCTTCCATCCCGGAATCTATGACC 600
DB 541 TCGAGAGAGGTTGAGGCTTATCAATGCTTCCATCCCGGAATCTATGACC 600
QY 601 CGGATTTCAATGCTGCAACCAAGAGAGGCTGCCAACAACGTCATCAAGGGCAAAAG 660
DB 601 CGGATTTCAATGCTGCAACCAAGAGAGGCTGCCAACAACGTCATCAAGGGCAAAAG 660
QY 661 AAGAGCAAGTTCAACAAATCATTAAGACATCAAGGCTTTAAAGAACCAACAAAGTGG 720
DB 661 AAGAGCAAGTTCAACAAATCATTAAGACATCAAGGCTTTAAAGAACCAACAAAGTGG 720
QY 721 ACAAGGTGTTGACTGTGGACTGCCAACAAGAGGTTACAGTAAATTTGTTGGGGCC 780
DB 721 ACAAGGTGTTGACTGTGGACTGCCAACAAGAGGTTACAGTAAATTTGTTGGGGCC 780
QY 781 TTAATGACACCATGAGAAATCTCTTGGCTGTGACAGAAATGAGGCTGAGATTCTC 840
DB 781 TTAATGACACCATGAGAAATCTCTTGGCTGTGACAGAAATGAGGCTGAGATTCTC 840
QY 841 CTTCACCTTTGATGTCATTTGCTTGTGTTATGAAAAATGTTCTTTCATTAATGAGACC 900
DB 841 CTTCACCTTTGATGTCATTTGCTTGTGTTATGAAAAATGTTCTTTCATTAATGAGACC 900
QY 901 CTGAGAAACCTTTTGTACACGAGGCTGATGATCTTGCAATCGGAGAAACATTTAATG 960
DB 901 CTGAGAAACCTTTTGTACACGAGGCTGATGATCTTGCAATCGGAGAAACATTTAATG 960
QY 961 GTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGATTTCTTG 1020
DB 961 GTGAGATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTGTTGATTTCTTG 1020

```

```

QY 1021 TGGGGGCTGTATCAAGCAACATCTATGACTGATTAACAACCATCTGGAAAAATGATG 1080
DB 1021 TGGGGGCTGTATCAAGCAACATCTATGACTGATTAACAACCATCTGGAAAAATGATG 1080
QY 1081 GATGATCTTTCCGCTCCACAACCTTTCCGTTCCAGGAAATCTTCCAGAGCAAGTTG 1140
DB 1081 GATGATCTTTCCGCTCCACAACCTTTCCGTTCCAGGAAATCTTCCAGAGCAAGTTG 1140
QY 1141 TTGATGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGAACATCCAGACATG 1200
DB 1141 TTGATGATATGCTCAACAGCAATGCCATCTCTATGAGCTGTGAACATCCAGACATG 1200
QY 1201 TTGTTGTTAATTAAGTATGCTTACGTAGGGAACAGAAAGAGGACATGATGATGACA 1260
DB 1201 TTGTTGTTAATTAAGTATGCTTACGTAGGGAACAGAAAGAGGACATGATGATGACA 1260
QY 1261 CTTCAAGATATTCATGAGGTGGAAGAGCAACATGTTTGCAACAACATCGAGAGATT 1320
DB 1261 CTTCAAGATATTCATGAGGTGGAAGAGCAACATGTTTGCAACAACATCGAGAGATT 1320
QY 1321 CCTCTTACCTGCTCTATTAATCTTGAATGCTGCTTCTTGTGAGCTCAGACTAGAA 1380
DB 1321 CCTCTTACCTGCTCTATTAATCTTGAATGCTGCTTCTTGTGAGCTCAGACTAGAA 1380
QY 1381 TCGAGTTTAAAGCTGAAAAATGAGGAAATTCACATCTCACCCAGTTGCTACCATCC 1440
DB 1381 TCGAGTTTAAAGCTGAAAAATGAGGAAATTCACATCTCACCCAGTTGCTACCATCC 1440
QY 1441 TCAGTACTCTCAACCAAGCTCCTCTGTGTTCAACCGGTTCAACAGTGTGAATGCAATTG 1500
DB 1441 TCAGTACTCTCAACCAAGCTCCTCTGTGTTCAACCGGTTCAACAGTGTGAATGCAATTG 1500
QY 1501 CAAGAGCGCTCAATGCTGAAAAATATAGAGGCTTGTGTGATGTTGCCCCAGAGA 1560
DB 1501 CAAGAGCGCTCAATGCTGAAAAATATAGAGGCTTGTGTGATGTTGCCCCAGAGA 1560
QY 1561 ATTAACAATGTTTCAGATCAAGTGAAGATGGAACCGAAGAAATATATGTTGGGGTGG 1620
DB 1561 ATTAACAATGTTTCAGATCAAGTGAAGATGGAACCGAAGAAATATATGTTGGGGTGG 1620
QY 1621 CCTAGCTGATGTTTATGTTATTAATATATGTTGCTTAAATTTTGCAGATGATTTGAA 1680
DB 1621 CCTAGCTGATGTTTATGTTATTAATATATGTTGCTTAAATTTTGCAGATGATTTGAA 1680
QY 1681 TGCATCAGCTTCAATTAATGCTTTAGAGCGGGCATATCTGTTACTAGAAACATGAATG 1740
DB 1681 TGCATCAGCTTCAATTAATGCTTTAGAGCGGGCATATCTGTTACTAGAAACATGAATG 1740
QY 1741 AATGATGATTAATTTTGTGT 1760
DB 1741 AATGATGATTAATTTTGTGT 1760

```

RESULT 3
 US-10-424-598-70167
 ; Sequence 70167, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5323)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 70167
 ; LENGTH: 1989
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_34372C.1
US-10-424-599-70167

Query Match 99.2%; Score 1745.8; DB 18; Length 1989;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1758; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1 CTCCTCTTATTCCTTTGTAATTCATTCATCTTAACTCTTGTGAAAAATATATGTC 60
DB 164 CTCCTCTTATTCCTTTGTAATTCATTCATCTTAACTCTTGTGAAAAATATATGTC 223
QY 61 TCGAATATTTAAGGTGAGTGTCTTAATGTAGATACCGGAGCTGAGATTCAGTCCG 120
DB 224 TCGAATATTTAAGGTGAGTGTCTTAATGTAGATACCGGAGCTGAGATTCAGTCCG 283
QY 121 TGTACACTACGAAACCAACCGAATCTTGTACAGAAACAGGAATGGCACTTACGTGA 180
DB 284 TGTACACTACGAAACCAACCGAATCTTGTACAGAAACAGGAATGGCACTTACGTGA 343
QY 181 TTGTCAAAACCAAAATCTGTCAATACGAATTTAAAAACAATCCATGTTCTTAATTAG 240
DB 344 TTGTCAAAACCAAAATCTGTCAATACGAATTTAAAAACAATCCATGTTCTTAATTAG 403
QY 241 GGGTAATGCTTGTGGGTT- GGGGTGGAACCAACGGCTCAACCCCTCAACGGTGTGTAAT 299
DB 404 GGGTAATGCTTGTGGGTTGGGGTGGGAACCAACGGCTCAACCCCTCAACGGTGTGTAAT 463
QY 300 GGTAAACGAGAGGAGATTCATAGGGCTAACAAAGAACAAATTCACAAAGCCATTTACT 359
DB 464 GGTAAACGAGAGGAGATTCATAGGGCTAACAAAGAACAAATTCACAAAGCCATTTACT 523
QY 360 GGCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGGAATCTAT 419
DB 524 GGCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGGAATCTAT 583
QY 420 GGGCAATTCAGAGCGCTGCTTCCATGTGTTAAACCTGACGACATGTGTTGGGGATGG 479
DB 584 GGGCAATTCAGAGCGCTGCTTCCATGTGTTAAACCTGACGACATGTGTTGGGGATGG 643
QY 480 GATATCAGCAATGATACTGGCTGATGCGCATGCGCAAGGGCAAGGTGTTTACATCAT 539
DB 644 GATATCAGCAATGATACTGGCTGATGCGCATGCGCAAGGGCAAGGTGTTTACATCAT 703
QY 540 TTGCAAGAGCAGTGAAGGCTTACATGGAATCATGCTTCCACTCCCGGAATCTATGAC 599
DB 704 TTGCAAGAGCAGTGAAGGCTTACATGGAATCATGCTTCCACTCCCGGAATCTATGAC 763
QY 600 CCGGATTTCAATGCTGCCCAACCAAGAGACGTGCCAACACATCATCAAGGGCAACAAAG 659
DB 764 CCGGATTTCAATGCTGCCCAACCAAGAGACGTGCCAACACATCATCAAGGGCAACAAAG 823
QY 660 CAAGAGCAGTTCACAAATCATCAAAAGCATCAAGAGCGTTTAAAGAACCCCAAAAGTG 719
DB 824 CAAGAGCAGTTCACAAATCATCAAAAGCATCAAGAGCGTTTAAAGAACCCCAAAAGTG 883
QY 720 GACAAGGTGTTGTAAGTGTGACTGCCAACACAGAGAGGTACATTAATTGGTGTGGGC 779
DB 884 GACAAGGTGTTGTAAGTGTGACTGCCAACACAGAGAGGTACATTAATTGGTGTGGGC 943
QY 780 GTTAAATGACACCAATGAGATCTTGGCTGCTGTGACAGAAATGAGGCTGAGATTTCT 839
DB 944 GTTAAATGACACCAATGAGATCTTGGCTGCTGTGACAGAAATGAGGCTGAGATTTCT 1003
QY 840 CTTTCCACCTTGTATGCCATGCTTGTATGAAATGTTCTTTTCAATTAATGGAAC 899
DB 1004 CTTTCCACCTTGTATGCCATGCTTGTGTATGAAATGTTCTTTTCAATTAATGGAAC 1063
QY 900 CCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCATGCGGAGAACACTTTGAT 959
DB 1064 CCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCATGCGGAGAACACTTTGAT 1123
QY 960 GGTGAGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTTGATTTCTT 1019
DB 1019 GGTGAGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTTGATTTCTT 1019
```

```

DB 1124 GGTGAGATGACTTCAAGAGTGTGACACCAAAATGAAATCTGTGTGTTGATTTCTT 1183
QY 1020 GTGGGGGCTGGTATCAAGCCCAACATCTATAGTTCATCAATCTCTGGAAACATGAT 1079
DB 1184 GTGGGGGCTGGTATCAAGCCCAACATCTATAGTTCATCAATCTCTGGAAACATGAT 1243
QY 1080 GGTATGAATCTTGTGGCTCCACAACTTCCGTTCCAAAGAAATCTCCAAAGCAACGTT 1139
DB 1244 GGTATGAATCTTGTGGCTCCACAACTTCCGTTCCAAAGAAATCTCCAAAGCAACGTT 1303
QY 1140 GTTATGATATGTGTCACAGCAATGCCATCTCTATAGGCGCTGGTGAACATCCAGACAT 1199
DB 1304 GTTATGATATGTGTCACAGCAATGCCATCTCTATAGGCGCTGGTGAACATCCAGACAT 1363
QY 1200 GTTGTGTTATTAAGTATGTGCTCTTAAGGAGGACAGCAAGAGCCATGATGATAC 1259
DB 1364 GTTGTGTTATTAAGTATGTGCTCTTAAGGAGGACAGCAAGAGCCATGATGATAC 1423
QY 1260 ACTTCAGAGATATTCATGGGTGGAAGAGACCAATGTTTTGACACACATGCGAGAT 1319
DB 1424 ACTTCAGAGATATTCATGGGTGGAAGAGACCAATGTTTTGACACACATGCGAGAT 1483
QY 1320 TCCCTCTTACGCTGCTCTTATATCTTGAACCTTGCTCTTCTTGTGACTGACACTGAGA 1379
DB 1484 TCCCTCTTACGCTGCTCTTATATCTTGAACCTTGCTCTTCTTGTGACTGACACTGAGA 1543
QY 1380 ATCGAGTTTAAAGCTGAAATGAGGGAATTTCCACTCATTCACCCAGTTGCTACATC 1439
DB 1544 ATCGAGTTTAAAGCTGAAATGAGGGAATTTCCACTCATTCACCCAGTTGCTACATC 1603
QY 1440 CTCAGCTACCTCACCAAGGCTCTCTGTGTTCCACCGGGTACACAGTGTGATGATG 1499
DB 1604 CTCAGCTACCTCACCAAGGCTCTCTGTGTTCCACCGGGTACACAGTGTGATGATG 1663
QY 1500 TCAAGACGCTGCAATCTGGAACAACTTAATGAGGCTGTGTTGATGTTGGCCCCAGAG 1559
DB 1664 TCAAGACGCTGCAATCTGGAACAACTTAATGAGGCTGTGTTGATGTTGGCCCCAGAG 1723
QY 1560 AATAACATGATTTCTGAGTACAGTGAAGTGGGACCGAAGAAATATATAGTTGGGGTA 1619
DB 1724 AATAACATGATTTCTGAGTACAGTGAAGTGGGACCGAAGAAATATATAGTTGGGGTA 1783
QY 1620 GCGTAGCTGAATGTTTATGTTAATATATGTTTGTCTTAATTTTGAAGTGAATGA 1679
DB 1784 GCGTAGCTGAATGTTTATGTTAATATATGTTTGTCTTAATTTTGAAGTGAATGA 1843
QY 1680 ATGCATCAGCTTCATTAATGCTTTAGACGGGGCATATTTGTTTACTAGAAACATGAT 1739
DB 1844 ATGCATCAGCTTCATTAATGCTTTAGACGGGGCATATTTGTTTACTAGAAACATGAT 1903
QY 1740 GAATGTAGTATTAATTTGTGT 1760
DB 1904 GAATGTAGTATTAATTTGTGT 1924
```

```

RESULT 4
US-10-025-003-9
; Sequence 9, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
```


NUMBER OF SEQ ID NOS: 16
 SOFTWARE: Microsoft Office 97
 SEQ ID NO: 9
 LENGTH: 1533
 TYPE: DNA
 ORGANISM: Glycine max
 US-10-025-003-9

Query Match 87.1%; Score 1533; DB 14; Length 1533;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

54 ATGTTATGAGAAATTTAAAGTTGAGTGTCTTAATGTAAGTACCGAGATGAGATT 113
 1 ATGTTATGAGAAATTTAAAGTTGAGTGTCTTAATGTAAGTACCGAGATGAGATT 60
 114 CAGTCCGTATCAACATACGAAACACCGAATTGTTACAGAGAACAGAAATGACCTAT 173
 61 CAGTCCGTATCAACATACGAAACACCGAATTGTTACAGAGAACAGAAATGACCTAT 120
 174 CAGTGGATGTCAAACCCCAATGTCCTCAATACGAAATTTAAACCAATCCATGTTCT 233
 121 CAGTGGATGTCAAACCCCAATGTCCTCAATACGAAATTTAAACCAATCCATGTTCT 180
 234 AATTAGGGGTAATGCTGTGGGTGGGTAAGAAACAGGCTCAACCGTCAACGGTGGT 293
 181 AATTAGGGGTAATGCTGTGGGTGGGTAAGAAACAGGCTCAACCGTCAACGGTGGT 240
 294 GTTATGCTAACCGAGGATTTTCATGAGGCTACAAAGAACAGAAATTCACCAAGCCAT 353
 241 GTTATGCTAACCGAGGATTTTCATGAGGCTACAAAGAACAGAAATTCACCAAGCCAT 300
 354 TACTTTGGCTCCCTCAACCCCAAGCTCACTATCCAGTTGGTCTTTCCAGGAGAGAA 413
 301 TACTTTGGCTCCCTCAACCCCAAGCTCACTATCCAGTTGGTCTTTCCAGGAGAGAA 360
 414 ATCTATGCCCCAATTGAAGGCTGCTTCAATGGTTAACCTGACCAATGTTGTTGGG 473
 361 ATCTATGCCCCAATTGAAGGCTGCTTCAATGGTTAACCTGACCAATGTTGTTGGG 420
 474 GATGAGATATCAGCAATGAACTGCTGATGCAATGGCCAGGCAAAAGGTTTGAC 533
 421 GATGAGATATCAGCAATGAACTGCTGATGCAATGGCCAGGCAAAAGGTTTGAC 480
 534 ATCGATTTCAAGAGAGATTGAGGCTTACATGAATTCATGCTTCCCTCCGGAATC 593
 481 ATCGATTTCAAGAGAGATTGAGGCTTACATGAATTCATGCTTCCCTCCGGAATC 540
 594 TATGACCCGGATTTTCTGCTGCCAACAGAGAGAGCTGCCAACAGCTCATCAAGGCG 653
 541 TATGACCCGGATTTTCTGCTGCCAACAGAGAGAGCTGCCAACAGCTCATCAAGGCG 600
 654 ACAAGCAAGAGAGATTCAACAAATCATCAAGAGATCAAGGCTTTTAAAGAGGACCC 713
 601 ACAAGCAAGAGAGATTCAACAAATCATCAAGAGATCAAGGCTTTTAAAGAGGACCC 660
 714 AAAGTGACAGAGTGTGTACTGCTGACTGCCAACAGAGAGATCAAGTAATTTGGTT 773
 661 AAAGTGACAGAGTGTGTACTGCTGACTGCCAACAGAGAGATCAAGTAATTTGGTT 720
 774 GTGGGCTTAATACCAATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 833
 721 GTGGGCTTAATACCAATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
 834 ATTTCTCTTCCACTTTGATGCTGCTGTTGATGAAATTTCTCTTCAATTAAT 893
 781 ATTTCTCTTCCACTTTGATGCTGCTGTTGATGAAATTTCTCTTCAATTAAT 840
 894 GGAAGCCCTCAGAAACTTTTGTACAGAGGCTGATTTGCTTCAGGAGAACT 953
 841 GGAAGCCCTCAGAAACTTTTGTACAGAGGCTGATTTGCTTCAGGAGAACT 900
 954 TTGATTGGTGAAGATCTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGTTGAT 1013

DB 901 TTGATTGGTGAAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTTGAT 960
 1014 TTCTTTGGGGGCTGTGATACAGCCCAATCTATAGTCAATACCATCTGGAAAC 1073
 DB 961 TTCTTTGGGGGCTGTGATACAGCCCAATCTATAGTCAATACCATCTGGAAAC 1020
 1074 AATGATGTAATATCTTTGGCTCCACAACTTTCCGTTCCAAAGAAATCTCCAAAGAC 1133
 DB 1021 AATGATGTAATATCTTTGGCTCCACAACTTTCCGTTCCAAAGAAATCTCCAAAGAC 1080
 1134 AACGTTGTGATGATATGATGTCACAGCAATGTCATCTCTATAGAGCTGTGAACATCA 1193
 DB 1081 AACGTTGTGATGATATGATGTCACAGCAATGTCATCTCTATAGAGCTGTGAACATCA 1140
 1194 GACCATGTTGTTATTAATGATGTCCTTAAGAGGAGACAGAGACCAATGAT 1253
 DB 1141 GACCATGTTGTTATTAATGATGTCCTTAAGAGGAGACAGAGACCAATGAT 1200
 1254 GAGTACACTTCAGAGATATTCATGGGTGAAAGAGACCAATGTTTTCACACATAGC 1313
 DB 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAGACCAATGTTTTCACACATAGC 1260
 1314 GAGATTCCTCTTAGCTGCTCTTAATCTTGAATGCTGCTTCTTGTAGCTCAGC 1373
 DB 1261 GAGATTCCTCTTAGCTGCTCTTAATCTTGAATGCTGCTTCTTGTAGCTCAGC 1320
 1374 ACTAGATGAGTTTAAAGCTGAAAAATGAGGAAAAATTCACATTCACACCAATGCT 1433
 DB 1321 ACTAGATGAGTTTAAAGCTGAAAAATGAGGAAAAATTCACATTCACACCAATGCT 1380
 1434 ACCATCTCAGTACTTACCTACCAAGGCTCTCTGTGTTCCACGGGTGACACCATGCTGAT 1493
 DB 1381 ACCATCTCAGTACTTACCTACCAAGGCTCTCTGTGTTCCACGGGTGACACCATGCTGAT 1440
 1494 GCATTGTCAAAGCAGGTCGATGCTGAAAAATGAGAGGCTTGTGTTGATTTGCC 1553
 DB 1441 GCATTGTCAAAGCAGGTCGATGCTGAAAAATGAGAGGCTTGTGTTGATTTGCC 1500
 1554 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1586
 DB 1501 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1533

RESULT 5
 US-10-718-952-9
 Sequence 9, Application US/10718952
 Publication No. US20040128713A1
 GENERAL INFORMATION:
 APPLICANT: Hiltz, William
 APPLICANT: Sebastian, Scott
 APPLICANT: Grace, John
 APPLICANT: Streitz, Leon
 TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAPEINOS
 TITLE OF INVENTION: SACCARIDES AND PHYTIC ACID
 FILE REFERENCE: BB-1077-C
 CURRENT APPLICATION NUMBER: US/10/718,952
 CURRENT FILING DATE: 2003-11-21
 PRIOR APPLICATION NUMBER: 08/835,751
 PRIOR FILING DATE: APRIL 8, 1997
 PRIOR APPLICATION NUMBER: PCT/US98/06822
 PRIOR FILING DATE: APRIL 7, 1998
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: Microsoft Office 97
 SEQ ID NO: 9
 LENGTH: 1533
 TYPE: DNA
 ORGANISM: Glycine max
 US-10-718-952-9
 Query Match 87.1%; Score 1533; DB 19; Length 1533;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 ATGTTTCATCGAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGAT 113
DB 1 ATGTTTCATCGAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGAT 60
QY 114 CAGTCGCGTGTACACTAGAAACCAACGAACTTGTTCAGAGAAAGAAATGGACCTAT 173
DB 61 CAGTCGCGTGTACACTAGAAACCAACGAACTTGTTCAGAGAAAGAAATGGACCTAT 120
QY 174 CAGTGGATTTGTCAAACTTGTCAAAATGTCAAAATTTAAACCAATTCATGTCTCT 233
DB 121 CAGTGGATTTGTCAAACTTGTCAAAATGTCAAAATTTAAACCAATTCATGTCTCT 180
QY 234 AAATTTAGGGGTATAGCTTGTGGGTGGGGTGAACCAACGCTCAACCTCAACCGGTGT 293
DB 181 AAATTTAGGGGTATAGCTTGTGGGTGGGGTGAACCAACGCTCAACCTCAACCGGTGT 240
QY 234 GTTATTTGCTAACCGAGAGGCAATTTATGAGGCTAACAAAGAACAAATTCACCAACCAAT 353
DB 241 GTTATTTGCTAACCGAGAGGCAATTTATGAGGCTAACAAAGAACAAATTCACCAACCAAT 300
QY 354 TACTTTGGTCCCTCAACCAACCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCT 413
DB 301 TACTTTGGTCCCTCAACCAACCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCT 360
QY 414 ATCTATGCCCCCAATTAAGAGCCTGCTTCAATGTGTTAACTTGAAGCAATGTGTTGGG 473
DB 361 ATCTATGCCCCCAATTAAGAGCCTGCTTCAATGTGTTAACTTGAAGCAATGTGTTGGG 420
QY 474 GGATGGGATATCAGCAACATGAACCTGCTGATGTCATGCGAGGCAAAAGGTGTTGAC 533
DB 421 GGATGGGATATCAGCAACATGAACCTGCTGATGTCATGCGAGGCAAAAGGTGTTGAC 480
QY 534 ATCGATTTGCAAGAGAGTGGAGGCTTACATGGAATCATGCTTCCATCTCCCGGAATC 593
DB 481 ATCGATTTGCAAGAGAGTGGAGGCTTACATGGAATCATGCTTCCATCTCCCGGAATC 540
QY 594 TATGACCCGGATTTTCAATGCTGCTCAACCAAGAGAGCGTCCCAACAGCTCATCAAGGAC 653
DB 541 TATGACCCGGATTTTCAATGCTGCTCAACCAAGAGAGCGTCCCAACAGCTCATCAAGGAC 600
QY 654 ACAAGCAAGAGCAAGTTCACCAATCATCAAGATCAAGAGCGTTCATTAAGAGAGCACC 713
DB 601 ACAAGCAAGAGCAAGTTCACCAATCATCAAGATCAAGAGCGTTCATTAAGAGAGCACC 660
QY 714 AAAGTGGACAAAGTGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGT 773
DB 661 AAAGTGGACAAAGTGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGT 720
QY 774 GTGGGCTTAATGACACATGAGGAATCTCTGGGCTGTGTGTAAGTGTGTAAGTGTGTAAG 833
DB 721 GTGGGCTTAATGACACATGAGGAATCTCTGGGCTGTGTGTAAGTGTGTAAGTGTGTAAG 780
QY 834 ATTTCTCTTCCACCTTGTATGCAATGCTTGTGTATGAAATTTCTTTCATTAAT 893
DB 781 ATTTCTCTTCCACCTTGTATGCAATGCTTGTGTATGAAATTTCTTTCATTAAT 840
QY 894 GGAAGCCCTCAAGACCTTTTGTACAGAGGCTGATTTGCTTCCAGAGAAACAT 953
DB 841 GGAAGCCCTCAAGACCTTTTGTACAGAGGCTGATTTGCTTCCAGAGAAACAT 900
QY 954 TTGATTTGGGAGATGATCTTCAAGAGTGTGACCAACCAATGAAATCTGTGTTGTTGAT 1013
DB 901 TTGATTTGGGAGATGATCTTCAAGAGTGTGACCAACCAATGAAATCTGTGTTGTTGAT 960
QY 1014 TTCTTTGGGGGCTGTATCAAGCCACATCTATAGTCAAGTTACCAACATCTGGGAAAC 1073
DB 961 TTCTTTGGGGGCTGTATCAAGCCACATCTATAGTCAAGTTACCAACATCTGGGAAAC 1020
QY 1074 AATGATGATGATTTTTCGCTCCAAATCTTTCGTTCCAGAGAAATCTTCCAGAGAC 1133
DB 1021 AATGATGATGATTTTTCGCTCCAAATCTTTCGTTCCAGAGAAATCTTCCAGAGAC 1080

QY 1134 AACGTTGTGATGATATGCTCAACAGCAATGCCATCTCTATGAGCCTGTGTAACATCCA 1193
DB 1081 AACGTTGTGATGATATGCTCAACAGCAATGCCATCTCTATGAGCCTGTGTAACATCCA 1140
QY 1194 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAAGGGGACACCAAGAGCCATGGAT 1253
DB 1141 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAAGGGGACACCAAGAGCCATGGAT 1200
QY 1254 GAGTACCTTCAGAGATATATCATGAGGTGGAAGAGACACATTTGTTTTCACAAACATATG 1313
DB 1201 GAGTACCTTCAGAGATATATCATGAGGTGGAAGAGACACATTTGTTTTCACAAACATATG 1260
QY 1314 GAGATTCCTCTTATGCTGCTCTATTAATCTTGAATGTTGCTTCTTGTGAGCTCAGC 1373
DB 1261 GAGATTCCTCTTATGCTGCTCTATTAATCTTGAATGTTGCTTCTTGTGAGCTCAGC 1320
QY 1374 ACTAGATTCGAGTTTAAAGCTGAAATTAAGGAAATTTCCATCTCATTTCCACCCAGTTGCT 1433
DB 1321 ACTAGATTCGAGTTTAAAGCTGAAATTAAGGAAATTTCCATCTCATTTCCACCCAGTTGCT 1380
QY 1434 ACCATCTCAGCTACCTCACCAGGCTCCTGTTCCACCGGGTACACAGTGTGAT 1493
DB 1381 ACCATCTCAGCTACCTCACCAGGCTCCTGTTCCACCGGGTACACAGTGTGAT 1440
QY 1494 GCATTTGCAAGACAGCGTGCATATGCTGGAATAACATTAATGAGGGCTTGTGTTGATGGCC 1553
DB 1441 GCATTTGCAAGACAGCGTGCATATGCTGGAATAACATTAATGAGGGCTTGTGTTGATGGCC 1500
QY 1554 CCAGGAATTAACATGATTTCTGAGTACAAAGTGA 1586
DB 1501 CCAGGAATTAACATGATTTCTGAGTACAAAGTGA 1533

RESULT 6

US-10-025-003-5
; Sequence 5, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-5

Query Match 87.0%; Score 1531.4; DB 14; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 54 ATGTTTCATCGAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGAT 113
DB 1 ATGTTTCATCGAATTTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGAT 60
QY 114 CAGTCGCGTGTACACTAGAAACCAACGAACTTGTTCAGAGAAAGAAATGGACCTAT 173
DB 61 CAGTCGCGTGTACACTAGAAACCAACGAACTTGTTCAGAGAAAGAAATGGACCTAT 120
QY 174 CAGTGGATTTGTCAAACTTGTCAAAATGTCAAAATTTAAACCAATTCATGTCTCT 233

```

Db 121 CAGTGGATTGCAAAACCAAAATCTGTCAAAATGAAATTTAAACCAACATTCATGTTCT 180
Qy 234 AATTTAGGGGTAATGTTGGGTTGGGGTGAATAAAGCGCTCAACCTTCAACGGTGGT 293
Db 181 AATTTAGGGGTAATGTTGGGTTGGGGTGAATAAAGCGCTCAACCTTCAACGGTGGT 240
Qy 294 GTTATTGCTAACGAGAGGCAATTTATGAGGCTTACAAAGGACAAAGATTTCAACAGCAAT 353
Db 241 GTTATTGCTAACGAGAGGCAATTTATGAGGCTTACAAAGGACAAAGATTTCAACAGCAAT 300
Qy 354 TACTTGGGCTCCCTCAACCAAGCTGAGTATCGAGTTGGGTCCTTCCAGGGAGAGAA 413
Db 301 TACTTGGGCTCCCTCAACCAAGCTGAGTATCGAGTTGGGTCCTTCCAGGGAGAGAA 360
Qy 414 ATCTATGCCCCATTCAGAGCGCTGCTTCCATAGTTAACTTGAACATTTGTTGGG 473
Db 361 ATCTATGCCCCATTCAGAGCGCTGCTTCCATAGTTAACTTGAACATTTGTTGGG 420
Qy 474 GGATGGGATATGAGCAATGAACCTGGCTGATGCGATGGCCAGGGCAAGGTTGTGAC 533
Db 421 GGATGGGATATGAGCAATGAACCTGGCTGATGCGATGGCCAGGGCAAGGTTGTGAC 480
Qy 534 ATCGATTTGCAAGAGCAGTTGAGGCTTTACATGGAATTCATGCTTCCCTCCGGAATC 593
Db 481 ATCGATTTGCAAGAGCAGTTGAGGCTTTACATGGAATTCATGCTTCCCTCCGGAATC 540
Qy 594 TATGACCCGGAATTTCTTGTCTGCAACCAAGAGAGCGTCCAAACAGCTCATCAAGGCT 653
Db 541 TATGACCCGGAATTTCTTGTCTGCAACCAAGAGAGCGTCCAAACAGCTCATCAAGGCT 600
Qy 654 ACAAAGCAAGAGCAAGTTCAACAAATCATCAAGAGCATCAAGGCTTTAAGGAAGCAGC 713
Db 601 ACAAAGCAAGAGCAAGTTCAACAAATCATCAAGAGCATCAAGGCTTTAAGGAAGCAGC 660
Qy 714 AAAGTGAACAAGGTTGTAAGTGTGAGCTGCAACAGAGAGGTAAGTAAATTTGTT 773
Db 661 AAAGTGAACAAGGTTGTAAGTGTGAGCTGCAACAGAGAGGTAAGTAAATTTGTT 720
Qy 774 GTGGGCTTTAATGACCACTGAGAAATCTTGGCTGTGTGAGCAAGAAATGAGCTGAG 833
Db 721 GTGGGCTTTAATGACCACTGAGAAATCTTGGCTGTGTGAGCAAGAAATGAGCTGAG 780
Qy 834 ATTTCTCTTCCACCTGTGATGCAATGCTGTGTGATGGAAGAAATGCTTCTTCAATAT 893
Db 781 ATTTCTCTTCCACCTGTGATGCAATGCTGTGTGATGGAAGAAATGCTTCTTCAATAT 840
Qy 894 GGAAGCCCTCAGAGACATTTGTACAGGCTGATGATCTTGCCATCGAGAAACT 953
Db 841 GGAAGCCCTCAGAGACATTTGTACAGGCTGATGATCTTGCCATCGAGAAACT 900
Qy 954 TTGATTTGGTGAAGTACTTCAAGAGTGTGACCAAAATGAATCTGTGTTGGTGAAT 1013
Db 901 TTGATTTGGTGAAGTACTTCAAGAGTGTGACCAAAATGAATCTGTGTTGGTGAAT 960
Qy 1014 TTCCTTGGGGGGCTGTATCAAGCAACATCTATATAGTAACTTCAACATCTTGGGAAC 1073
Db 961 TTCCTTGGGGGGCTGTATCAAGCAACATCTATATAGTAACTTCAACATCTTGGGAAC 1020
Qy 1074 AATGATGATGAATCTTTCGCTCCAAACCTTTCGCTTCCAGGAATCTTCAAGAGC 1133
Db 1021 AATGATGATGAATCTTTCGCTCCAAACCTTTCGCTTCCAGGAATCTTCAAGAGC 1080
Qy 1134 AAAGTTGATGATATGATGTCACAGCAATGCTATCTTATAGCTGTGTAACATCA 1193
Db 1081 AAAGTTGATGATATGATGTCACAGCAATGCTATCTTATAGCTGTGTAACATCA 1140
Qy 1194 GACCAATGTTGTTAATGATGCTTACGTAAGGGAAGCAAGAGAGCCATGAT 1253
Db 1141 GACCAATGTTGTTAATGATGCTTACGTAAGGGAAGCAAGAGAGCCATGAT 1200
Qy 1254 GAGTACACTTCAAGATATTCATGAGTGAAGAGACCAATGTTTGAACAACATGC 1313
Db 1201 GAGTACACTTCAAGATATTCATGAGTGAAGAGACCAATGTTTGAACAACATGC 1260

```

```

Qy 1314 GAGATTCCTCTGTAGCTGCTCTATATCTTGAATTTGGTCTTCTTGTGAGCTCAGC 1373
Db 1261 GAGGATTCCTCTGTAGCTGCTCTATATCTTGAATTTGGTCTTCTTGTGAGCTCAGC 1320
Qy 1374 ACTAGATGAGTTTAAAGCTGAAATGAGGGAATTTCCACTATTTCAACCAAGTGGCT 1433
Db 1321 ACTAGATGAGTTTAAAGCTGAAATGAGGGAATTTCCACTATTTCAACCAAGTGGCT 1380
Qy 1434 ACCATCTCAGCTTACCTCAACAGGCTCTGTGTTCCACGGGTAACACAGTGTGAAT 1493
Db 1381 ACCATCTCAGCTTACCTCAACAGGCTCTGTGTTCCACGGGTAACACAGTGTGAAT 1440
Qy 1494 GCAATGTCAAACAGCTGCAATGCTGGAATAATATATAGGCTTGTGTTGGATGGCC 1553
Db 1441 GCAATGTCAAACAGCTGCAATGCTGGAATAATATATAGGCTTGTGTTGGATGGCC 1500
Qy 1554 CCAGAGAAATATCATGATTTCTGAGTACAAAGTGA 1586
Db 1501 CCAGAGAAATATCATGATTTCTGAGTACAAAGTGA 1533

RESULT 7
US-10-718-952-5
; Sequence 5, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; PRIORITY FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-5

Query Match 87.0%; Score 1531.4; DB 19; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 54 ATGTTCAATCGAATTTTAAAGTTGAGTGTCTATATGAAATGATACCGAGACTGAGATT 113
Db 1 ATGTTCAATCGAATTTTAAAGTTGAGTGTCTATATGAAATGATACCGAGACTGAGATT 60
Qy 114 CAGTCCGTGTAACAATAAGCAACCAACCAACTTGTTCACGAGAACAGAAATGGCACTAT 173
Db 61 CAGTCCGTGTAACAATAAGCAACCAACCAACTTGTTCACGAGAACAGAAATGGCACTAT 120
Qy 174 CAGTGGATTTGCAAAACCAAAATCTGTCAAAATGCAATTTAAACCAACATCATCTTCT 233
Db 121 CAGTGGATTTGCAAAACCAAAATCTGTCAAAATGCAATTTAAACCAACATCATCTTCT 180
Qy 234 AAATTAAGGGGTAATGCTTGGGGTTGGGGTGAAGCAACAGGCTCAACCTCAACGGTGT 293
Db 181 AAATTAAGGGGTAATGCTTGGGGTTGGGGTGAAGCAACAGGCTCAACCTCAACGGTGT 240
Qy 294 GTTATTGCTAACGAGAGGCAATTTATGAGGCTTACAAAGGACAAAGATTTCAACAGCAAT 353
Db 241 GTTATTGCTAACGAGAGGCAATTTATGAGGCTTACAAAGGACAAAGATTTCAACAGCAAT 300
Qy 354 TACTTGGCTCCCTCACCAGGCTCAGCTATCGAGTTGGGTCCTTCCAGGGAGAGAA 413

```

```
Db 301 TACTTGGCTCCCTCAACCAAGCCCTCAGCTATCCAGTGGGTCTTCCAGGAGAGAA 360
Qy 414 ATCTATGCCCATTCGAAGAGCTGCTCCATAGTTAAACCTGACGACATTTGTTGGG 473
Db 361 ATCTATGCCCATTCGAAGAGCTGCTCCATAGTTAAACCTGACGACATTTGTTGGG 420
Qy 474 GGATGGATATCAGCAATGAAGCTGGCTGATGCGCATGGCCAAAGGTGTTTAC 533
Db 421 GGATGGATATCAGCAATGAAGCTGGCTGATGCGCATGGCCAAAGGTGTTTAC 480
Qy 534 ATCGATTTGCAAGACAGTGTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAATC 593
Db 481 ATCGATTTGCAAGACAGTGTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAATC 540
Qy 594 TATGACCCGGATTTTATGCTGCTGCAACCAAGAGAGCGGCAACAGTCAATCAAGGAC 653
Db 541 TATGACCCGGATTTTATGCTGCTGCAACCAAGAGAGCGGCAACAGTCAATCAAGGAC 600
Qy 654 ACAAGCAAGACAGTTCACCAATCATCAAGACATCAAGCGCTTTAAGGAAGCCAC 713
Db 601 ACAAGCAAGACAGTTCACCAATCATCAAGACATCAAGCGCTTTAAGGAAGCCAC 660
Qy 714 AAAGTGACAAAGTGTGTTGATCTGAGACTGCCAACACAGAGGTACATTAATTGGTT 773
Db 661 AAAGTGACAAAGTGTGTTGATCTGAGACTGCCAACACAGAGGTACATTAATTGGTT 720
Qy 774 GTGGGCTTTAATGACACCATGAGAGATCTTTGGGCTGTTGAGAGAAATGAGGCTAG 833
Db 721 GTGGGCTTTAATGACACCATGAGAGATCTTTGGGCTGTTGAGAGAAATGAGGCTAG 780
Qy 834 ATTTCCTCTCCACCTTGTATGCCATTTGCTTGTATGAAATGTTCTTTCTATAT 893
Db 781 ATTTCCTCTCCACCTTGTATGCCATTTGCTTGTATGAAATGTTCTTTCTATAT 840
Qy 894 GGAAGCCCTCAGAACATTTTGTACAGGCTGATGATCTTGCCATGGCGAGAACT 953
Db 841 GGAAGCCCTCAGAACATTTTGTACAGGCTGATGATCTTGCCATGGCGAGAACT 900
Qy 954 TTGATTTGGAGATGATCTTCAAGAGTGTGACACCAAAATGAATCTGTGTTGTTAT 1013
Db 901 TTGATTTGGAGATGATCTTCAAGAGTGTGACACCAAAATGAATCTGTGTTGTTAT 960
Qy 1014 TTCTTGTGGGGGCTGTATCAAGCCAACTATAGTCACTTACCAACCATCTGGAAAC 1073
Db 961 TTCTTGTGGGGGCTGTATCAAGCCAACTATAGTCACTTACCAACCATCTGGAAAC 1020
Qy 1074 AATGATGATGAATCTTTCGGCTCACAACATTTCCGTTCCAAAGAAATCTCCAGAGC 1133
Db 1021 AATGATGATGAATCTTTCGGCTCACAACATTTCCGTTCCAAAGAAATCTCCAGAGC 1080
Qy 1134 AACGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
Db 1081 AACGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Qy 1194 GACCAATGTTGTTATTAAGTATGTCCTTAAGAGGAGCAGCAAGAGCCATGAT 1253
Db 1141 GACCAATGTTGTTATTAAGTATGTCCTTAAGAGGAGCAGCAATAGAGCCATGAT 1200
Qy 1254 GAGTACATTCAGAGATATTCATGGGTGAAAAGAGACCATTTGTTGCAACACATGC 1313
Db 1201 GAGTACATTCAGAGATATTCATGGGTGAAAAGAGACCATTTGTTGCAACACATGC 1260
Qy 1314 GAGGATTCCTCTTACGCTGCTCTATTAATCTTGAATGCTGCTCTTGTGAGCTAGC 1373
Db 1261 GAGGATTCCTCTTACGCTGCTCTATTAATCTTGAATGCTGCTCTTGTGAGCTAGC 1320
Qy 1374 ACTAGAAATGAAATTAAGCTGAAAATGAGGAAAATTCATCTCAATCCAGTGGCT 1433
Db 1321 ACTAGAAATGAAATTAAGCTGAAAATGAGGAAAATTCATCTCAATCCAGTGGCT 1380
Qy 1434 ACCATCTCAGTACCTCAAGAGGCTCTGTGTTCCACCGGGTACACCAAGTGTGAT 1493
```

```
Db 1381 ACCATCTCAGTACCTCAAGAGGCTCTGTGTTCCACCGGGTACACCAAGTGTGAT 1440
Qy 1494 GCATTTGCAAGACAGCGTGCATATCTGAAAAATTAAGAGGCTTGTGATGATGGC 1553
Db 1441 GCATTTGCAAGACAGCGTGCATATCTGAAAAATTAAGAGGCTTGTGATGATGGC 1500
Qy 1554 CCAAGAAATTAATGATTTCTGAGTACAAATGA 1586
Db 1501 CCAAGAAATTAATGATTTCTGAGTACAAATGA 1533

RESULT 8
US-10-025-003-13
; Sequence 13, Application US/10025003
; Publication No. US20030074685A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streitz, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/025,003
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-025-003-13

Query Match 83.3%; Score 1465.8; DB 14; Length 1533;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1491; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 54 ATGTTTCATCGAATTTTAAGTGTGATGTCCTTAATGGAATGATACACCGAGATCGAGATT 113
Db 1 ATGTTTCATCGAATTTTAAGTGTGATGTCCTTAATGGAATGATACACCGAGATCGAGATT 60
Qy 114 CAGTCCGATGCAACTACGAACCAACCACTGTTCCAGAACAGAAATGGAACCTAT 173
Db 61 CAGTCCGATGCAACTACGAACCAACCACTGTTCCAGAACAGAAATGGAACCTAT 120
Qy 174 CAGTGGATTTGCAAAACCAATCTGTCAAAATAGAAATTTAAACCAACATCATGTTCTT 233
Db 121 CAGTGGATTTGCAAAACCAATCTGTCAAAATAGAAATTTAAACCAACATCATGTTCTT 180
Qy 234 AAATTAAGGGGTAATGCTTGTGGGTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGG 293
Db 181 AAATTAAGGGGTAATGCTTGTGGGTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGG 240
Qy 294 GTTATTTGTAACGAGAGGCAATTCATGGGCTACAAAGAGCAAGATTCACCAAGCAAT 353
Db 241 GTTATTTGTAACGAGAGGCAATTCATGGGCTACAAAGAGCAAGATTCACCAAGCAAT 300
Qy 354 TACTTTGGCTCCTCACCAGGCTCAGCTATCCGATTTGGGTCCTTCAGGAGAGAA 413
Db 301 TACTTTGGCTCCTCACCAGGCTCAGCTATTCGAGTTGGATCTTTCAGGAGAGAA 360
Qy 414 ATCTATGCCCATTCGAAGAGCTGCTCCATAGTTAAACCTGACGACATTTGTTGGG 473
Db 361 ATCTATGCCCATTCGAAGAGCTGCTCCATAGTTAAACCTGACGACATTTGTTGGG 420
Qy 474 GGATGGATATCAGCAATGAAGCTGGCTGATGCGCATGGCCAAAGGTGTTTAC 533
Db 421 GGATGGATATCAGCAATGAAGCTGGCTGATGCGCATGGCCAAAGGTGTTTAC 480
```

QY 534 ATGATTTGACAGAGAGTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAATC 593
 DB 481 ATCGATTTGACAGAGAGTTGAGGCTTACATGAAATCCATGCTTCCACTCCCGGAATC 540
 QY 594 TATGACCCGATTTTCTGCTGCCAACCAAGAGAGCTGCCAACACGTCATCAAGGCG 653
 DB 541 TACGACCCGATTTTCTGCTGCCAACCAAGAGAGCTGCCAACACGTCATCAAGGCG 600
 QY 654 ACAAGACAGAGAGTTCAACAATCATCAAGAGATCAAGGCGTTTAAAGAGGACAC 713
 DB 601 ACAAGACAGAGAGTTCAACAATCATCAAGAGATCAAGGCGTTTAAAGAGGACAC 660
 QY 714 AAAGTGACAGAGTGTGTACTGTGAGCTGCCAACAGAGAGTACATATTTGTT 773
 DB 661 AAAGTGACAGAGTGTGTACTGTGAGCTGCCAACAGAGAGTACATATTTGTT 720
 QY 774 GTGGGCTTAAATGACCAATGAGAAATCTTGTGCTGTGAGACAGAAATGAGGCTGAG 833
 DB 721 GTAGGCTTAAATGACCAATGAGAAATCTTGTGCTGTGAGACAGAAATGAGGCTGAG 780
 QY 834 ATTTCTCTTCCACCTGTATGACCATTTGCTGTATGAGAAATGTTCTTTCATTAAT 893
 DB 781 ATTTCTCTTCCACCTGTATGACCATTTGCTGTATGAGAAATGTTCTTTCATTAAT 840
 QY 894 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCATCCGAGAACACT 953
 DB 841 GGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCATCCGAGAACACT 900
 QY 954 TTGATTTGTGAGAGTACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGTTGAT 1013
 DB 901 TTGATTTGTGAGAGTACTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGTTGAT 960
 QY 1014 TTCTTTGTGAGGAGTGTATCAAGGCAATCATATGTGATGACATCAACATCTGGGAAAC 1073
 DB 961 TTCTTTGTGAGGAGTGTATCAAGGCAATCATATGTGATGACATCAACATCTGGGAAAC 1020
 QY 1074 AATGATGTGATGATCTTGTGCTCCACAACTTTCCGTCCAAAGAAATCTCCAAAGAC 1133
 DB 1021 AATGATGTGATGATCTTGTGCTCCACAACTTTCCGTCCAAAGAAATCTCCAAAGAC 1080
 QY 1134 AATGTTGTGATGATGTGACAGCAATGCCATCTGTATGAGCTGTGTAACATCCA 1193
 DB 1081 AATGTTGTGATGATGTGACAGCAATGCCATCTGTATGAGCTGTGTAACATCCA 1140
 QY 1194 GACCAATGTTGTATTAAGTATGCTTACATGAGGAGACAGCAAGAGACCATGAT 1253
 DB 1141 GACCAATGTTGTATTAAGTATGCTTACATGAGGAGATGACAAAGAGACCATGAT 1200
 QY 1254 GAGTACACTTCAAGATATTTCAATGAGTGGTGAAGAGACCATGTTTTCACACATGC 1313
 DB 1201 GAGTACACTTCAAGATATTTCAATGAGTGGTGAAGAGACCATGTTTTCACACATGC 1260
 QY 1314 GAGATTCCTCTTACAGTCTCTTATTTATCTTGAATGCTGCTTCTTTCAGTACAC 1373
 DB 1261 GAGATTCCTCTTACAGTCTCTTATTTATCTTGAATGCTGCTTCTTTCAGTACAC 1320
 QY 1374 ACTAATATGAGTATTAAGTGAAGTGAAGAAATTCATCTATCCATCCACATGTTGT 1433
 DB 1321 ACTAATATGAGTATTAAGTGAAGTGAAGAAATTCATCTATCCATCCACATGTTGT 1380
 QY 1434 ACCATCTCAGTACCTCACAAGGCTCTGTGTTCCACCGGATACACAGTGTGTAAT 1493
 DB 1381 ACCATCTCAGTACCTCACAAGGCTCTGTGTTCCACCGGATACACAGTGTGTAAT 1440
 QY 1494 GATTTGTCAAAAGCAGTGAATGCTGAGAAACATPATGAGGCTTGTGTTGATTTGCTC 1553
 DB 1441 GATTTGTCAAAAGCAGTGAATGCTGAGAAACATPATGAGGCTTGTGTTGATTTGCTC 1500
 QY 1554 CCAGAGATTAAGATATCTGAGTACAAATGTA 1586
 DB 1501 CCAGAGATTAAGATATCTGAGTACAAATGTA 1533

RESULT 9
 US-10-025-003-15
 ; Sequence 15. Application US/10025003
 ; Publication No. US20030074685A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hiltz, William
 ; APPLICANT: Sebastian, Scott
 ; APPLICANT: Grace, John
 ; APPLICANT: Streitz, Leon
 ; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
 ; FILE REFERENCE: BB-1077-C
 ; CURRENT APPLICATION NUMBER: US/10/025,003
 ; PRIOR FILING DATE: 2002-05-07
 ; PRIOR APPLICATION NUMBER: 08/835,751
 ; PRIOR FILING DATE: APRIL 8, 1997
 ; PRIOR APPLICATION NUMBER: PCT/US98/06822
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO. 15
 ; LENGTH: 1533
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; US-10-025-003-15
 Query Match 83.3%; Score 1465.8; DB 14; Length 1533;
 Best Local Similarity 97.3%; Pred. No. 0;
 Matches 1491; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
 QY 54 ATGTCATGAGAAATTTTAAAGTGTAGTGTCCATATGTAAGTACACGAGACTGAGATT 113
 DB 1 ATGTCATGAGAAATTTTAAAGTGTAGTGTCCATATGTAAGTACACGAGACTGAGATT 60
 QY 114 CAGTCGTGTACACTACGAAACCAACGAACTTGTTCAGAGAACAGAAATGGCACTAT 173
 DB 61 CAGTCGTGTACACTACGAAACCAACGAACTTGTTCAGAGAACAGAAATGGCACTAT 120
 QY 174 CAGTGTATGTCAAAACCCAAATCTGTCAAAATTCGAATTTAAACCAACATTCATGTTCT 233
 DB 121 CAGTGTATGTCAAAACCCAAATCTGTCAAAATTCGAATTTAAACCAACATTCATGTTCT 180
 QY 234 AAATTAAGGGGTAAATGCTGTGAGTGTGGGGTGAACCAAGGCTCAACCTGACCGGTGAT 293
 DB 181 AAATTAAGGGGTAAATGCTGTGAGTGTGGGGTGAACCAAGGCTCAACCTGACCGGTGAT 240
 QY 294 GTTATTTGCTAACCGAGAGGCAATTCATGAGGCTTACAAAGAGACAAATTCACAAAGCAAT 353
 DB 241 GTTATTTGCTAACCGAGAGGCAATTCATGAGGCTTACAAAGAGACAAATTCACAAAGCAAT 300
 QY 354 TACTTTGGCTCCCTCAACCCAGAGCTGAGTATCCGAGTTGGGCTCTTCCAGGAGAGAA 413
 DB 301 TACTTTGGCTCCCTCAACCCAGAGCTGAGTATCCGAGTTGGGCTCTTCCAGGAGAGAA 360
 QY 414 ATCTATGCCCAATTCAGAGGCTGCTTCCAAAGTGTAAACCTCGAAGACATGTTGTTGGG 473
 DB 361 ATCTATGCCCAATTCAGAGGCTGCTTCCAAAGTGTAAACCTCGAAGACATGTTGTTGGG 420
 QY 474 GGATGGATATTCAGAGACATGAACCTGCTGATGTCATGAGCCAGAGGCAAGAGTGTGTAAC 533
 DB 421 GGATGGATATTCAGAGACATGAACCTGCTGATGTCATGAGCCAGAGGCAAGAGTGTGTAAC 480
 QY 534 ATGATTTGCAAGACAGTTGAGGCTTATCATGAAATCATGCTTCCACTCCCGGAATC 593
 DB 481 ATGATTTGCAAGACAGTTGAGGCTTATCATGAAATCATGCTTCCACTCCCGGAATC 540
 QY 594 TATGACCCGGAATTTATGTTGCTGCAACCAAGAGAGCGGTCCAAAGTCAATCAAGGCG 653
 DB 541 TATGACCCGGAATTTATGTTGCTGCAACCAAGAGAGCGGTCCAAAGTCAATCAAGGCG 600
 QY 654 ACAAGCAAGAGCAAGTTCAACAATCATCAAGAGATCAAGGCGTTTAAAGAGCCACAC 713
 DB 601 ACAAGCAAGAGCAAGTTCAACAATCATCAAGAGATCAAGGCGTTTAAAGAGCCACAC 660

QY 714 AAAGTGAACAAGTGTGTACTGTGAGCTGCGCAACACAGAGAGGTACAGTAATTTGGT 773
DB 661 AAAGTGAACAAGTGTGTGTCTGTGTGAGTGTGCAACACAGAGAGGTATATGCAATTTGGT 720
QY 774 GTGGGCTTAATGACACCAATGAGAAATCTCTTGAGTGTGTGAGAGAAATGAGGCTGAG 833
DB 721 GTAGGCTTAATGACACCAATGAGAAATCTCTTGAGTGTGTGAGAGAAATGAGGCTGAG 780
QY 834 ATTTCTCTTCCACCTTGTATGAGCAATGCTGTGTATGAGAAATGCTTCAATTAAT 893
DB 781 ATTTCTCTTCCACCTTGTATGAGCAATGCTGTGTATGAGAAATGCTTCAATTAAT 840
QY 894 GGAAGCCCTCAGAAACCTTTGTACAGAGGCTGATGATCTTGCCATGCGAGAGAACT 953
DB 841 GGAAGCCCTCAGAAACCTTTGTACAGAGGCTGATGATCTTGCCATGCGAGAGAACT 900
QY 954 TTGATTTGGTGAAGATGACTTCAAGAGTGTGTGAGCAACCAAAATGAAATCTGTGTGAT 1013
DB 901 TTGATTTGGTGAAGATGACTTCAAGAGTGTGTGAGCAACCAAAATGAAATCTGTGTGAT 960
QY 1014 TTCTTTGTGGGGCTGTGTATCAAGCCAACTCTATATGTCAGTTACATCATTTGGGAAAC 1073
DB 961 TTTCTTTGTGGGGCTGTGTATCAAGCCAACTCTATATGTCAGTTACATCATTTGGGAAAC 1020
QY 1074 AATGATGGTATGATCTTTGCGGCTCCAAACCTTCCGTTCCAGAGAAATCTCCAAGAGC 1133
DB 1021 AATGATGGTATGATCTTTGCGGCTCCAAACCTTCCGTTCCAGAGAAATCTCCAAGAGC 1080
QY 1134 AACGTTGTATGATATGTCACCAAGCAATGCTCTTATGAGCTGTGTGAGCAATCCA 1193
DB 1081 AACGTTGTATGATATGTCACCAAGCAATGCTCTTATGAGCTGTGTGAGCAATCCA 1140
QY 1194 GACCATGTTGTGTATTAATGATGTGCTTACGTATGAGGAGCAAGCAAGAGCCATGAT 1253
DB 1141 GACCATGTTGTGTATTAATGATGTGCTTACGTATGAGGAGCAAGCAAGAGCCATGAT 1200
QY 1254 GAGTACACTTCAAGATATTCATGAGGTGAGAAAGAGACCATGTTTGTGCAACAACATSC 1313
DB 1201 GAGTACACTTCAAGATATTCATGAGGTGAGAAAGAGACCATGTTTGTGCAACAACATSC 1260
QY 1314 GAGATTTCCCTCTTACGCTCTCTATTAATCTTGAATGCTCTTGTGCTGTAGCTCAGC 1373
DB 1261 GAGATTTCCCTCTTACGCTCTCTATTAATCTTGAATGCTCTTGTGCTGTAGCTCAGC 1320
QY 1374 ACTAGATGAGTTTAACTGAAATGAGGAGAAATTCACACTCATTCACCAAGTTGCT 1433
DB 1321 ACTAGATGAGTTTAACTGAAATGAGGAGAAATTCACACTCATTCACCAAGTTGCT 1380
QY 1434 ACCATCTCAGCTACCTCAACAGAGCTCTGCTGTCCACCGGCTACCAAGTGTGAT 1493
DB 1381 ACCATCTCAGCTATGTACCAAGGCTCTCTGTGTTCCACCGGCTACCAAGTGTGAT 1440
QY 1494 GCATTGTCAAGCAGCTGCAATGTGAGAAACATTAATGAGGCTGTGTGTGATTTGGCC 1553
DB 1441 GCATTGTCAAGCAGCTGCAATGTGAGAAACATTAATGAGGCTGTGTGTGATTTGGCC 1500
QY 1554 CCAAGAGATTAACATGATTTCTCGAGTCAAGTGA 1586
DB 1501 CCAAGAGATTAACATGATTTCTCGAGTCAAGTGA 1533

RESULT 10
US-10-718-952-13
; Sequence 13, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; TITLE OF INVENTION: SACHARIDES AND PHYTIC ACID

FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718.952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835.751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-13

Query Match 83.3%; Score 1465.8; DB 19; Length 1533;
Beet Local Similarity 97.3%; Pred. No. 0;
Matches 1491; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 54 ATGTCATCGAGAAATTTTAAGTTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATT 113
DB 1 ATGTCATCGAGAAATTTTAAGTTGAGTGTCTTAATGTGAAGTACACCGAGACTGAGATT 60
QY 114 CAGTCCGTGTACACTACGAAACCAACCGAATTTGTTCAAGAGACAGAAATGGAACCTTAT 173
DB 61 CAGTCCGTGTACACTACGAAACCAACCGAATTTGTTCAAGAGACAGAAATGGAACCTTAT 120
QY 174 CAGTGGATTTGCAACCAAAATCTGTCAAAATACGAATTTAAACCAACATCATGTTCT 233
DB 121 CAGTGGATTTGCAACCAAAATCTGTCAAAATACGAATTTAAACCAACATCATGTTCT 180
QY 234 AAATTAGGGGTATATGCTGTGTGGGTGAGAAACAAACGCTCAACCTCACCGGTGT 293
DB 181 AAATTAGGGGTATATGCTGTGTGGGTGAGAAACAAACGCTCAACCTCACCGGTGT 240
QY 294 GTTATTTGCTAACCGAGAGGCTTTTCAATGAGCTTAAAGAGACAAATTTCAACCAAT 353
DB 241 GTTATTTGCTAACCGAGAGGCTTTTCAATGAGCTTAAAGAGACAAATTTCAACCAAT 300
QY 354 TACTTGGCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGGAA 413
DB 301 TACTTGGCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGGAA 360
QY 414 ATCTATGCCCATTTAAAGAGCTGCTTCAATGTTAAACCTGACGACATGTGTTGGG 473
DB 361 ATCTATGCCCATTTAAAGAGCTGCTTCAATGTTAAACCTGACGACATGTGTTGGG 420
QY 474 GGATGGATATACGACAACTGAGCTGTGATGCTGAGGCAAGGCAAGGCTTTGATC 533
DB 421 GGATGGATATACGACAACTGAGCTGTGATGCTGAGGCAAGGCAAGGCTTTGATC 480
QY 534 ATCGATTTGCAAGACAGGTGAGGCTTAAACATGAATCATGCTTCCACTCCCGGAATC 593
DB 481 ATCGATTTGCAAGACAGGTGAGGCTTAAACATGAATCATGCTTCCACTCCCGGAATC 540
QY 594 TATGACCCGGAATTTGATGCTGCTCAACAGAGAGCGTGTCCCAACAGCTCATCAAGGCG 653
DB 541 TATGACCCGGAATTTGATGCTGCTCAACAGAGAGCGTGTCCCAACAGCTCATCAAGGCG 600
QY 654 ACAAGCAAGACAGATTTCAACAAATCATCAAAAGCATCAAGGCTTTAAGAGACCAAC 713
DB 601 ACAAGCAAGACAGATTTCAACAAATCATCAAAAGCATCAAGGCTTTAAGAGACCAAC 660
QY 714 AAAGTGAACAAGTGTGTACTGTGAGCTGCGCAACACAGAGAGGTACAGTAATTTGGT 773
DB 661 AAAGTGAACAAGTGTGTGTCTGTGTGAGTGTGCAACACAGAGAGGTATATGCAATTTGGT 720
QY 774 GTGGGCTTAATGACACCAATGAGAAATCTCTTGAGTGTGTGAGAGAAATGAGGCTGAG 833
DB 721 GTAGGCTTAATGACACCAATGAGAAATCTCTTGAGTGTGTGAGAGAAATGAGGCTGAG 780
QY 834 ATTTCTCTTCCACCTTGTATGAGCAATGCTGTGTATGAGAAATGCTTCAATTAAT 893

```
Db      781 ATTTCTCTTCCACCTTGATGCCATTCGTGTGTGAGTGAATAATGTTCTTTCAATTAAT 840
Qy      894 GGAAGCCCTTCAGAACTTTTGTACCAAGGCTGATTTGATTTTCCATTCGGAGAACT 953
Db      841 GGAAGCCCTTCAGAACTTTTGTACCAAGGCTGATTTGATTTTCCATTCGGAGAACT 900
Qy      954 TTGATTTGGTGAATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 1013
Db      901 TTGATTTGGTGAATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 960
Qy      1014 TTCTTTTGGGGGCTGTATCAAGCAACATCTATAGTCAGTTACCAACCATCTGGAAAC 1073
Db      961 TTTCTTTGGGGGCTGTATCAAGCAACATCTATAGTTACCAACCATCTGGAAAC 1020
Qy      1074 AATGATGATGATCTTTTGGGCTCCACAACTTTCCGTTCCAGGAATCTTCAAGAGC 1133
Db      1021 AATGATGATGATCTTTTGGGCTCCACAACTTTCCGTTCCAGGAATCTTCAAGAGC 1080
Qy      1134 AACGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
Db      1081 AACGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Qy      1194 GACCATGTTGTTGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
Db      1141 GACCATGTTGTTGTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Qy      1254 GAGTACACTTCAGAGATATTCATGGGTGGAAGAGACCAATGTTTTCACACATGTC 1313
Db      1201 GAGTACACTTCAGAGATATTCATGGGTGGAAGAGACCAATGTTTTCACACATGTC 1260
Qy      1314 GAGGATTTCCCTTCTACTGCTGCTCTATTTATCTTGAGATGCTGCTCTGCTGAGCTCAG 1373
Db      1261 GAGGATTTCCCTTCTACTGCTGCTCTATTTATCTTGAGATGCTGCTCTGCTGAGCTCAG 1320
Qy      1374 ACTAGATGAGATTTTAAAGCTGAAATGAGGAAATTCACATCTATTCACACCATGTCCT 1433
Db      1321 ACTAGATGAGATTTTAAAGCTGAAATGAGGAAATTCACATCTATTCACACCATGTCCT 1380
Qy      1434 ACCATCTCACTGCTCACTCAACCAAGGCTCTCTGTTTCCACCGGTAACACAGTGTGAAT 1493
Db      1381 ACCATCTCACTGCTCACTCAACCAAGGCTCTCTGTTTCCACCGGTAACACAGTGTGAAT 1440
Qy      1494 GCATTGTCAAGAGGCTGCAATGCTGGAATAACATATGAGGCTGTGTGATTTGGCC 1553
Db      1441 GCATTGTCAAGAGGCTGCAATGCTGGAATAACATATGAGGCTGTGTGATTTGGCC 1500
Qy      1554 CCAGAGATTAACATGATCTCGAGTACAAGTGA 1586
Db      1501 CCAGAGATTAACATGATCTCGAGTACAAGTGA 1533

RESULT 11
US-10-718-952-15
; Sequence 15, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hiltz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streit, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718, 952
; PRIOR FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 9, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1533
```

```
; TYPE: DNA
; ORGANISM: Glycine max
; US-10-718-952-15
Query Match      83.3%; Score 1465.8; DB 19; Length 1533;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1491; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy      54  AAGTTCATGAGAAATTTTAAAGTTGAGTGTCTTATGTGAGTACACCGAGCTGAGATT 113
Db      1  AAGTTCATGAGAAATTTTAAAGTTGAGTGTCTTATGTGAGTACACCGAGCTGAGATT 60
Qy      114  CAGTCGTGTACAACTACGAAACCAACCGAATCTGTTTCAAGAAACGAAATGACCTAT 173
Db      61  CAGTCGTGTACAACTACGAAACCAACCGAATCTGTTTCAAGAAACGAAATGACCTAT 120
Qy      174  CAGTGAATGTCAAAACCAATCTGTCAAAATGCAATTTAAACCAACATCTGATTTCT 233
Db      121  CAGTGAATGTCAAAACCAATCTGTCAAAATGCAATTTAAACCAACATCTGATTTCT 180
Qy      234  AAATTAGGGGTAAATGCTGTGGGTGGGGTGAACCAACGCTCAACCTCAACCGGTGT 293
Db      181  AAATTAGGGGTATGCTGTGGGTGGGGTGAACCAACGCTCAACCTCAACCGGTGT 240
Qy      294  GTTATTTCTTACCGAGAGGCTTTTCAATGGGCTTACAAAGGAAAGATTTCAAGCCAT 353
Db      241  GTTATTTCTTACAGAGAGGCTTTTCAATGGGCTTACAAAGGAAAGATTTCAAGCCAT 300
Qy      354  TACTTTGGCTCCCTCAACCAACCTCAGCTATCCGATGTTGGTCTTCCAGGAGAGAA 413
Db      301  TACTTTGGCTCCCTCAACCAACCTCAGCTATCCGATGTTGGTCTTCCAGGAGAGAA 360
Qy      414  ATCTATGCCCCATTCAGAGCTGCTTCCATGTTTAAACCTGACGACATGTGTTGGG 473
Db      361  ATCTATGCCCCATTCAGAGCTGCTTCCATGTTTAAACCTGACGACATGTGTTGGG 420
Qy      474  GGAATGGATATGAGAACTGAACCTGGCTGATGTCATGGCCAGGCAAAAGGTGTTGAC 533
Db      421  GGAATGGATATGAGAACTGAACCTGGCTGATGTCATGGCCAGGCAAAAGGTGTTGAC 480
Qy      534  ATCGATTTTGCAGAAAGCTGAGGCTTTACATGGAATCATCTTCCACTCCCGGAATC 593
Db      481  ATCGATTTTGCAGAAAGCTGAGGCTTTACATGGAATCATCTTCCACTCCCGGAATC 540
Qy      594  TATGACCCGGATTTTATGCTGTCACCAAGAGAGCGTCCCAACACCTCATCAAGGCT 653
Db      541  TATGACCCGGATTTTATGCTGTCACCAAGAGAGCGTCCCAACACCTCATCAAGGCT 600
Qy      654  ACAAAAGCAAGAGCAATTCATCAAAATCATCAAGAGCGTCTTAAAGAACCTAC 713
Db      601  ACAAAAGCAAGAGCAATTCATCAAAATCATCAAGAGCGTCTTAAAGAACCTAC 660
Qy      714  AAAGTGAACAAGGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773
Db      661  AAAGTGAACAAGGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy      774  GTGGGCTTTAATGACACCATGAGAAATCTTGGCTGTGTGAGACGAATAGGCTGAG 833
Db      721  GTGGGCTTTAATGACACCATGAGAAATCTTGGCTGTGTGAGACGAATAGGCTGAG 780
Qy      834  ATTTTCTCTTCCACTTGTATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 893
Db      781  ATTTTCTCTTCCACTTGTATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy      894  GGAAGCCCTTCAGAACTTTTGTACCAAGGCTGATTTGATTTTCCATTCGGAGAACT 953
Db      841  GGAAGCCCTTCAGAACTTTTGTACCAAGGCTGATTTGATTTTCCATTCGGAGAACT 900
Qy      954  TTGATTTGGTGAATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 1013
Db      901  TTGATTTGGTGAATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 960
Qy      1014  TTCTTTTGGGGGCTGTATCAAGCAACATCTATAGTCAGTTACCAACCATCTGGAAAC 1073
```


Db	961	TTCTCTGGGGGCTGGTATCAGCCAACTCTATAGTATGTTACAACTCTGGGAAAC	1020
Qy	1074	AATGATGGTAAAGATCTTTCGGCTCCAAACCTTTCGTTCCAGGAATCTCCAAAGC	1133
Db	1021	AATATGTTATGAAATCTCTGGCTCCAAACCTTTCGGCTCCAAAGAAATCTCCAAAGC	1080
Qy	1134	AACGTTGTATGATATGTGTCAACAGCATGCCATCTCTATAGCCTGGTGAACATCCA	1193
Db	1081	AACGTTGTATGATATGTGTCAACAGCATGCCATCTCTATAGCCTGGTGAACATCCC	1140
Qy	1194	GACCATGTGTGTATTAATTAAGTATGTGCCTTACGTAGGGGACAGCAAGAGCCATGGAT	1253
Db	1141	GACCATGTGTGTATTAATTAAGTATGTGCCTTACGTAGGGGATAGCAAGAGCCATGGAT	1200
Qy	1254	GAGTACACTTCAGAGATATTCATGGGTGGAAAGAGACACATTTGTTTTGCACACATGTC	1313
Db	1201	GAGTACACTTCAGAGATATTCATGGGTGGAAAGAGACACATTTGTTTTGCACACATGT	1260
Qy	1314	GAGCATTCCTCTTACGTCTCTCTATTAATCTTGGACCTTGTCTTGTCTGAGCTAGC	1373
Db	1261	GAGCATTCCTCTTACGTCTCTCTATTAATCTTGGACCTTGTCTTGTCTGAGCTAGC	1320
Qy	1374	ACTAGAAATCGAGTTTAAAGCTGAAATAGAGGGAATTTCCACTCATTCACCCAGTTGCT	1433
Db	1321	ACTAGAAATCGAGTTTAAAGCTGAAATAGAGGGAATTTCCACTCATTCACCCAGTTGCT	1380
Qy	1434	ACCATTCCTCAGCTTACCTCACCAGGCTCTCTGTGTTCCACCGGGTACACAGTGTGAAT	1493
Db	1381	ACCATTCCTCAGCTTACCTCACCAGGCTCTCTGTGTTCCACCGGGTACACAGTGTGAAT	1440
Qy	1494	GCATTGTCAAGACGCGTGCATATGTGGAAAAACATAATGAGGCTGTGTGTGATTTGGCC	1553
Db	1441	GCATTGTCAAGACGCGTGCATATGTGGAAAAACATAATGAGGCTGTGTGTGATTTGGCC	1500
Qy	1554	CCAGAGAATAACATGATTCCTCGAGTACAAAGTGA	1586
Db	1501	CCAGAGAATAACATGATTCCTCGAGTACAAAGTGA	1533
RESULT 12			
US-10-025-003-11			
: Sequence 11, Application US/10025003			
: Publication No. US20030074685A1			
: GENERAL INFORMATION:			
: APPLICANT: Hitz, William			
: APPLICANT: Sebastian, Scott			
: APPLICANT: Grace, John			
: APPLICANT: Streif, Leon			
: TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE			
: TITLE OF INVENTION: SACCARIDES AND PHYTIC ACID			
: FILE REFERENCE: BB-1077-C			
: CURRENT APPLICATION NUMBER: US/10/025.003			
: CURRENT FILING DATE: 2002-05-07			
: PRIOR APPLICATION NUMBER: 08/835,751			
: PRIOR FILING DATE: APRIL 8, 1997			
: PRIOR APPLICATION NUMBER: PCT/US98/06822			
: PRIOR FILING DATE: APRIL 7, 1998			
: NUMBER OF SEQ ID NOS: 16			
: SOFTWARE: Microsoft Office 97			
: SEQ ID NO 11			
: LENGTH: 1533			
: TYPE: DNA			
: ORGANISM: Glycine max			
US-10-025-003-11			

Query Match	83.2%	Score 1464.2;	DB 14;	Length 1533;
Best Local Similarity	97.2%;	Pred. No. 0;		
Matches 1450; Conservative	0;	Mismatches 43;	Indels 0;	Gaps 0;

```

QY      54  ATGTCATCGAATTTTAAGGTGAGTCTCTAATGTAAGTACCCGAGACTGAGATT  11
        |||||
Db      1  ATGTCATCGAATTTTAAGGTGAGAGTCTCTAATGTAAGTACCCGAGACTGAGATT  60

```

QY	114	CAGTCCGCTGTAACAATAACGAACCAACCGAATTGTTACAGAAACAGAAATGCACTTAT	173
Db	61	CAGTCCGCTGTAACAATAACGAACCAACCGAATTGTTACAGAAACAGAAATGCACTTAT	120
QY	174	CAGTGAATGTCACAACCCCAATCTGTCAAAATTTGATAAATCAACATCTCATGTTCTCT	233
Db	121	CAGTGAATGTCACAACCCCAATCTGTCAACAATTGATAAATCAACATCTCATGTTCTCA	180
QY	234	AAATTTAGGGGTAAATGCTTGTGGGGTGGGGTGGAAACAACGACTTCAACCTTCAACGGTGT	293
Db	181	AAATTTAGGGGGTAAATGCTTGTGGGGTGGGGTGGAAACAACGAGCTTCAACCTTCAACGGTGT	240
QY	294	GTTATTTGCTAACGAGAGGGGCAATTTCAATGGGCTTCAAAAGGCAAGATTTCAACAAGCCAAAT	353
Db	241	GTTATTTGCTAACGAGAGGACATTTCAATGGGCTTCAAAAGGCAAGATTTCAACAAGCCAAAT	300
QY	354	TACTTTGGCTCCCTCACCCAAAGCCTCAGCTATCCGAGTTGGGCTCTTCCAGGAGAGAA	413
Db	301	TACTTTGGCTCCCTCACCCAAAGCCTCAGCTATTTGAGTTGGATCTCTTCCAGGAGAGAA	360
QY	414	ATCTATGCCCATTCACAGAGCCTGCTTCCAAATGGTTAAACCTGACGACATTTGTGTTGGG	473
Db	361	ATCTATGCCCATTCACAGAGCTGCTTCCAAATGGTTAACTTGACGACATTTGTGTTGGG	420
QY	474	GGATGGGATATCAGCAACATGAACCTTGCTGATGCCATGGCCAGAGGCAAAAGTGTTTGAC	533
Db	421	GGATGGGATATCAGCAACATGAACCTTGCTGATGCCATGGCCAGAGGCAAAAGGTTTGGAC	480
QY	534	ATCGATTTGCAAGAACAGTTGAGGCTTTACATGGAATTCATGCTTCCATCTCCCGGAATC	593
Db	481	ATCGATTTTGCAGAAAGCAGTTGAGGCTTTAAATGGAATTCATGTTTCCATCTCCCGGAATC	540
QY	594	TATGACCCGGATTTCAATGCTGCGCAACCAAGAGAGGCTGCAACAACGTCATCAAGGGC	653
Db	541	TACGACCCGGATTTCAATGCTGCTGCCAACCAAGAGAGGCTGCCAACAACGTCATTAAAGGC	600
QY	654	ACAAAGCAAGACAGCAAGTTCAACAATCATCAAGACATCAAGGCTTTAAAGAACCAACC	713
Db	601	ACAAAGCAAGACAGCAAGTTCAAGCAAAATCATCAAGACATCAAGGCTTTAAAGAACCAACC	660
QY	714	AAATGTGACAAAGTGTGTTACTCTGTGAACGCGCAACAACAAGAGGTACAGTAATTTGGTT	773
Db	661	AAAGTGAACAAGGTGTGTTCTCTGTGAACGCGCAACAACAAGAGGTATAGCAAAATTTGGTT	720
QY	774	GTGGGCTCTTAATGACACCATGAGAAATCTCTGGCTGCTGTGACAGAAAAATGAGGCTGAG	833
Db	721	GTAGGCTCTTAATGACACCATGAGAAATCTCTGGCTGCTGTGACAGAAAAATGAGGCTGAG	780
QY	834	ATTTTCTCCTTCCACCTTGTATGCAATGCTTGTGTGTATATGAAAAATGTTTCTTTCAATTAAT	893
Db	781	ATTTTCTCCTTCCACCTTGTATGCAATGCTGTGTGTATATGAAAAATGTTTCTTTCAATTAAT	840
QY	894	GGAAGCCCTTGAAACATTTTGTACCAAGGGCTATATATCTTGCTACGCGAGGAACACT	953
Db	841	GGAAGCCCTTGAAACACTTTTGTATCAAGGGCTATATATCTTGCTACGCGAGGAACACT	900
QY	954	TTGATTTGTGAGATGACTTCAAGAGTGTGACCACAAAATGAAATCTGTGTGTTGAT	1013
Db	901	TTGATTTGTGAGATGACTTCAAGAGTGTGACCACAAAATGAAATCTGTGTGTTGAT	960
QY	1014	TTCTCTTGTGGGGGTGATATCAACCCAAATCTATACAGTTTACAAACCTATCTGGGAAAC	1073
Db	961	TTCTCTTGTGGGGGTGATATCAACCCAAATCTATATGTTATGTTTCAACCAATCTGGGAAAC	1020
QY	1074	AATGATGGTATGAATCTTTGGGCTTCCAAACTTTCCGTTTCCAAGGAATCTCCAAAGC	1133
Db	1021	AATGATGGTATGAATCTCTGGGCTTCCAAACCTTCCGCTTCCAAGGAATCTCCAAAGC	1080
QY	1134	AACGTTGTTGATATGCTTCAACAGCAATGCCATCTCTATAGGCTCGGTGAACATCCA	1193
Db	1081	AACGTTGTTGATATGCTTCAACAGCAATGCCATCTCTATAGGCTCGGTGTGAACATCTCC	1140

Qy	1194	GACCATTGTTGTTGTTATTAAGTAATATGCTTACGTAGGGGACAGCAGACAGCCATGGAT	12533
Db	1141	GACCATTGTTGTTGTTATTAAGTAATATGCTTACGTAGGGGATAGCAAGAGCCATGGAT	12000
Qy	1254	GAGTACACTTCAGAGATATATCATGGGGTGGAAAGACACATTTGTTTTGCACAAACATGAC	13133
Db	1201	GAGTACACTTCAGAGATATATCATGGGGTGGAAAGACACCATTTGTTTTGCACAAACATGT	12660
Qy	1314	GAGGATTCCTCTTAGCTGCTCCTATTAATCTTGAATTGATCTTCTTGTGAGCTCAGC	13737
Db	1261	GAGGATTCCTCTTAGCTGCTCCTATTAATCTTGAATTGATCTTCTTGTGAGCTGAGC	13200
Qy	1374	ACTAGAAATCCAGGTTTAAAGCTGAAATAAGGGAAATTCACCTCATTTCCACCAGTTGCT	14333
Db	1321	ACTAGAAATCCAGGTTTAAAGCTGAAATAAGGGAAATTCACCTCATTTCCACCAGTTGCT	13800
Qy	1434	ACCATTCCCTCAGCTACCTCACCAGGCTCTCTGTTTCCACCGGGATCACCAGTGGTAAT	14933
Db	1381	ACCATTCCCTCAGCTACCTCACCAGGCTCTCTGTTTCCACCGGGATCACCAGTGGTAAT	14400
Qy	1494	GCATTGTCAAAAGCAGCGTGCATATGCTGAAAAACATATAGAGGGCTTGTTGGATTGGCC	15533
Db	1441	GCATTGTCAAAAGCAGCGTGCATATGCTGAAAAACATATAGAGGGCTTGTTGGATTGGCC	15000
Qy	1554	CCAGAGAAATACATGATTTCTGAGAGTACAGTGA	1586
Db	1501	CCAGAGAAATACATGATTTCTGAGAGTACAGTGA	1533

```

RESULT 13
US-10-718-952-11
; Sequence 11, Application US/10718952
; Publication No. US20040128713A1
; GENERAL INFORMATION:
; APPLICANT: Hitz, William
; APPLICANT: Sebastian, Scott
; APPLICANT: Grace, John
; APPLICANT: Streif, Leon
; TITLE OF INVENTION: SOYBEAN PLANT PRODUCING SEEDS WITH REDUCED LEVELS OF RAFFINOSE
; FILE REFERENCE: BB-1077-C
; CURRENT APPLICATION NUMBER: US/10/718,952
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: 08/835,751
; PRIOR FILING DATE: APRIL 8, 1997
; PRIOR APPLICATION NUMBER: PCT/US98/06822
; PRIOR FILING DATE: APRIL 7, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 11
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Glycine max
US-10-718-952-11

Query Match      83.2%; Score 1464.2; DB 19; Length 1533;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY      54      ATGTTTCATCGAGATTTTAAAGTTGAGTGTGCTTAATGTGAAGTACACCGAGACTGAGATT 113
Db      1      ATGTTTCATCGAGATTTTAAAGTTGAGAGTCTTAATGTGAAGTACACCGAGACTGAGATT 60

QY      114     CAGTCGCGTGTACAACTACGAAACCAACCGAACTTGTTACAGAGAAACAGAGAAATGGCACTTAT 173
Db      61     CAGTCGCGTGTACAACTACGAAACCAACCGAACTTGTTACAGAGAAACAGAGAAATGGCACTTAT 120

QY      174     CAGTGAGATTGTCAAAACCAAAATCTGTCAAAATAGCAATTTTAAACCAACCATCATGTTGCT 233
Db      121     CAGTGAGATTGTCAAAACCAAAATCCGTGCTACATACCAATTTTAAACCAACCAACCATGTTCCA 180

QY      234     AATTTAGGGGTAAATGCTTTGTGGGTGGGTGGAGAAACAAGGCTCAACCTTCACCGGTGCT 293

```

Db	181	AAATTGGGGGTGATGCTTGTGGGTTGGGGGTGAAAACAA	CGGCTCTACCTCA	CCGGTGT	240
QY	294	GTWATTTGTAACCCAGAGGGGATTTCAATGGGCTACAAAG	GAACAATTCA	CAAGCCAA	353
Db	241	GTATATGTCTAACAGAGGAGACATTTCAATGGGCTACAAAG	CAACAATTCA	CAAGCCAA	300
QY	354	TACTTTGGCTCCCTCAACCCAAAGCCTCAGCTATCCGAG	TTGGTCTCTTCA	GGGAGAGAA	413
Db	301	TACTTTGGCTCCCTCAACCCAAAGCCTCAGCTATCCGAG	TTGGTCTCTTCA	GGGAGAGAA	360
QY	414	ATCTATGGCCCAATTCAGAGGCTGCTTCCATATGCTTAA	CCCTGACGACATATG	TTTTGGG	473
Db	361	ATCTATGGCCCAATTCAGAGGCTGCTTCCATATGCTTAA	CCCTGACGACATATG	TTTTGGG	420
QY	474	GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATG	GGCCAGAGCAAGGTGTTGAC		533
Db	421	GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATG	GGCCAGAGCAAGGTGTTGAC		480
QY	534	ATCGATTTTGCAGAAAGCAGTTGAGGCTTTACATGAA	ATCATGCTTCCA	CTCCCGGAATC	593
Db	481	ATCGATTTTGCAGAAAGCAGTTGAGGCTTTACATGAA	ATCATGCTTCCA	CTCCCGGAATC	540
QY	554	TATGATCCCGGATTTCAATTTGCTGCCACCAAGAGAG	GGTGCCCAACAACTCAT	CAAGGGC	653
Db	541	TACGATCCCGGATTTCAATTTGCTGCCACCAAGAGAG	GGTGCCCAACAACTCAT	CAAGGGC	600
QY	654	ACAAAGCAAGAGCAAGTTCAACAAATCATCAAGACAT	CAAGGGGTTTAA	GAAGAGCCAC	713
Db	601	ACAAAGCAAGAGCAAGTTCAACAAATCATCAAGACAT	CAAGGGGTTTAA	GAAGAGCCAC	660
QY	714	AAAGTGCACAAAGTGTGTTGTA	CTGTGACCTGCCAA	CAGAGAGTACAGTAA	773
Db	661	AAAGTGCACAAAGTGTGTTGTA	CTGTGACCTGCCAA	CAGAGAGTATAGCAAT	720
QY	774	GTGGGCTTAAATGACACCATGSGAATCTCTTGGCTGT	CTGTGACAGAAATAGAG	CGCTGAG	833
Db	721	GTAGGCTCTTAATGACACCATGSGAATCTCTTGGCTGT	CTGTGACAGAAATAGAG	CGCTGAG	780
QY	834	ATTTCCTCTCCACCTTGATGACCATGCTGTGTTATG	AGAAATGTTCCCTTCA	TATTAAT	893
Db	781	ATTTCCTCTCCACCTTGATGACCATGCTGTGTTATG	AGAAATGTTCCCTTCA	TATTAAT	840
QY	894	GGAAGCCCTCAGAACACTTTTGTGACAGGGCTGATTG	ATCTTGGCATCGCAGAGAAC	ACT	953
Db	841	GGAAGCCCTCAGAACACTTTTGTGACAGGGCTGATTG	ATCTTGGCATCGCAGAGAAC	ACT	900
QY	954	TTGATTTGGTGAATGACCTTCAAGAGTGTGACACCAA	AATGAATCTGTGTGTTGAT		1013
Db	901	TTGATTTGGTGAATGACCTTCAAGAGTGTGACACCAA	AATGAATCTGTGTGTTGAT		960
QY	1014	TTTCTTTGGGGGGGTGGATATCAAGCCAAATCTATAT	CACTGTTTCAACCAATCTG	GGGAAAC	1073
Db	961	TTTCTTTGGGGGGGTGGATATCAAGCCAAATCTATAT	CACTGTTTCAACCAATCTG	GGGAAAC	1020
QY	1074	AATGATGGTATGAATCTTTCCGGCTCCACAAACTTT	CCGTTCCAGAGAAATCT	CCAAAGC	1133
Db	1021	AATGATGGTATGAATCTTTCCGGCTCCACAAACTTT	CCGTTCCAGAGAAATCT	CCAAAGC	1080
QY	1134	AACGTTGTGATGATATGCTTCAACAGCAATGCCAT	CTCTATAGGCTGTGAA	CATCCA	1193
Db	1081	AACGTTGTGATGATATGCTTCAACAGCAATGCCAT	CTCTATAGGCTGTGAA	CATCCC	1140
QY	1194	GACCATGTTGTTGTTATTAAGTATGTCCTTACGTAG	GGGATACAGAGCCAT	GGAT	1253
Db	1141	GACCATGTTGTTGTTATTAAGTATGTCCTTACGTAG	GGGATACAGAGCCAT	GGAT	1200
QY	1254	GAGTACACTTCAGAGATATTCATGGGTGAAAGAGCA	CAATTTGTTTGCA	CAACACATGC	1313
Db	1201	GAGTACACTTCAGAGATATTCATGGGTGAAAGAGCA	CAATTTGTTTGCA	CAACACATGT	1260
QY	1314	GAGGATTCCTCTTACGCTGCTCTTATATCTTGAC	CTTGGTCTTCTTCTTG	CTGAC	1373
Db	1261	GAGGATTCCTCTTACGCTGCTCTTATATCTTGAC	CTTGGTCTTCTTCTTG	CTGAC	1320

QY 1374 ACTAGATCGAGTTTAAAGCTGAAAAATGAGGAAAAATTCCTCATTCACCCAGTTGCT 1433
| | | | |
Db 1321 ACTAGATCGAGTTTAAAGCTGAAAAATGAGGAAAAATTCCTCATTCACCCAGTTGCT 1380
| | | | |
QY 1434 ACCATCTCAGCTACCTCCACCAAGGCTCTCTGTTCCACCGGGTACACCACTGTGTAAT 1493
| | | | |
Db 1381 ACCATCTCAGCTACCTCCACCAAGGCTCTCTGTTCCACCGGGTACACCACTGTGTAAT 1440
| | | | |
QY 1494 GATTTGTCAAAGCAGCTGCAATGCTTGAAACATTAAGAGGGCTTGCTTGGAATTTGGCT 1553
| | | | |
Db 1441 GATTTGTCAAAGCAGCTGCAATGCTTGAAACATTAAGAGGGCTTGCTTGGAATTTGGCT 1500
| | | | |
QY 1554 CCAGAGAAATACATGATTTCTCGAGTACCAAGTGA 1586
| | | | |
Db 1501 CCAGAGAAATACATGATTTCTCGAGTACCAAGTGA 1533
| | | | |

RESULT 14

US-10-424-599-12022
; Sequence 12022, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12022
; LENGTH: 2018
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110663C.1
US-10-424-599-12022

Query Match 73.7%; Score 1296.4; DB 18; Length 2018;

Best Local Similarity 88.6%; Pred. No. 0;

Matches 1405; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 47 AAAAAAATGTTCAATCGAGAAATTTAAAGTTAGTGTCTTAATGGAAGTACCGAGAC 106
| | | | |
Db 91 AAAAAAATGTTCAATCGAGAAATTTAAAGTTAGTGTCTTAATGGAAGTACCGAGAC 150
| | | | |
QY 107 TGAGATTCAGTCCGCTGTAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 166
| | | | |
Db 151 TGAGATTCAGTCCGCTGTAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 210
| | | | |
QY 167 CACTATACAGTGGATTTGCAAAACCAATCTGTCAAAATGCAATTTAAAAACAACATCCA 226
| | | | |
Db 211 CACTATACAGTGGATTTGCAAAACCAATCTGTCAAAATGCAATTTAAAAACAACATCCA 270
| | | | |
QY 227 TCTTCTTAATTAAGGGGTAAATGTTGGGTGGGGTGGAAACAACGGCTCAACCTCAC 286
| | | | |
Db 271 TCTTCTTAATTAAGGGGTAAATGTTGGGTGGGGTGGAAACAACGGCTCAACCTCAC 330
| | | | |
QY 287 CGATGCTTATTTGCTAAACGAGAGGGCATTTCAATGGCTACAAAGACAAAGATTTCAACA 346
| | | | |
Db 331 TGGTGTGTATTTGCTAAACGAGAGGGCATTTCAATGGCTACAAAGACAAAGATTTCAACA 390
| | | | |
QY 347 ACCCAATTAATTTGGCTCCCTCAACCAAGCCTCAAGTATCCAGTTGGGTCTTCACAGG 406
| | | | |
Db 391 ACCCAATTAATTTGGCTCCCTCAACCAAGCCTCAAGTATCCAGTTGGGTCTTCACAGG 450
| | | | |
QY 407 AAGAGAAATCTATGCCCCCTTCAAGAGCTGCTTCCAAATGTTAAACCTGAGACATTTGT 466
| | | | |
Db 451 GGAAGAAATATATGCTCAATTCAGAGCCTGCTTCCAAATGTTAAACCTGAGATGATTTGT 510
| | | | |

QY 467 GTTTGGGGGATGGGATATACGAACAATGAACCTGCTGATGCCATATGGCCAGGGCAAGGT 526
| | | | |
Db 511 GTTTGGGGGATGGGATATACGAACAATGAACCTGCTGATGCCATATGGCCAGGGCAAGGT 570
| | | | |
QY 527 GTTTGACATCGATTTTGCAGAAAGCAGTTGAGGCCCTTAATGGAATTCATGCTTTCCACTGCC 586
| | | | |
Db 571 GTTTGATATCGAACCTGCGAGAAACAGTTGAGGCCCTTAATGGAATTCATGCTTTCCACTGCC 630
| | | | |
QY 587 CGGAATCTATGACCCGGATTTTCATTTGCTGCAACCAAGAGAGGGCTGCAACCAAGTAT 646
| | | | |
Db 631 CGGAATCTATGACCCGGATTTTCATTTGCTGCAACCAAGAGAGGGCTGCAACCAAGTAT 690
| | | | |
QY 647 CAAGGGCAAAAGCAAGCAAGTTCACAAATCATCAAGACATCAAGCGTTTAAAGA 706
| | | | |
Db 691 CAAGGGCAAAAGCAAGCAAGTTCACAAATCATCAAGACATCAAGCGTTTAAAGA 750
| | | | |
QY 707 AGCCACCAAGTGAACAAGGTGTTGTAATGCTGGAATGCCCAACACAGAGAGTACAGTAA 766
| | | | |
Db 751 AAGCACTAAGGTTGACAAGGTGTTGTCCTGTGACAGCCCAACACAGAGAGTACAGCAA 810
| | | | |
QY 767 TTTGGTTGGGCTTAATGACACCATGAGAAATCTTTGGCTGCTGTCAGAGAAATGA 826
| | | | |
Db 811 GGTAGTTGTGGACTTAACGACACCATGAAACCTCTTTGCTTCTTGACAGAGAAATGA 870
| | | | |
QY 827 GGCTGAGATTTCTCCTTCACCTGTGATGCCATGCTTTGTGTAAGAAAATGTTCTTT 886
| | | | |
Db 871 GGCTGAGATTTCTCCTTCACCTGTGATGCCATGCTTTGTGTAAGAAAATGTTCTTT 930
| | | | |
QY 887 CATTAATGGAAGCCTTCAGAACCTTTGTATCCAGGGGTGAATTAATCTTGCATCCGAC 946
| | | | |
Db 931 CATCAATGGAAGCCTTCAGAACCTTTGTATCCAGGGGTGAATTAATCTTGCATCCGAC 990
| | | | |
QY 947 GAACACTTTGATGTTGGTGAAGATGACTTCAAGAGTGTGACAGCAAAATGAATCTGTGTT 1006
| | | | |
Db 991 GAATGATTTGATTTGAGAGAGATGACTTAAAGTGTGACAGCAAAATGAATGATGTT 1050
| | | | |
QY 1007 GGTGATTTCTTGTGGGGGCTGTGATCAAGCCCAATCTTAATGCTAGTTACAAACATCT 1066
| | | | |
Db 1051 GGTGATTTCTTGTGAGAGAGATGACTTAAAGTGTGACAGCAAAATGAATGATGTT 1110
| | | | |
QY 1067 GGGAAACAATGATGATTAATCTTTGGCTCCCAACCTTCCGTTCCAGAGAAATCTC 1126
| | | | |
Db 1111 GGGAAACAATGATGATTAATCTTTGGCTCCCAACCTTCCGTTCCAGAGAAATCTC 1170
| | | | |
QY 1127 CAAGAGCAACCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186
| | | | |
Db 1171 CAAGAGCAACCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1230
| | | | |
QY 1187 ACATCCAGACCATGTTGTTGTTAATTAAGTATGCTTAAAGTGAAGGAGACAGCAAGAGAC 1246
| | | | |
Db 1231 ACATCCAGACCATGTTGTTGTTAATTAAGTATGCTTAAAGTGAAGGAGACAGCAAGAGAC 1290
| | | | |
QY 1247 CATGATGATGATCACTTCAAGATTAATTAAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1306
| | | | |
Db 1291 CATGATGATGATCACTTCAAGATTAATTAAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1350
| | | | |
QY 1307 CACATGAGAGATTCCTCTTAAGCTGCTCTAATTAATCTTGAACCTGCTCTCTCTCTCTCTCTCTCT 1366
| | | | |
Db 1351 CACATGAGAGATTCCTCTTAAGCTGCTCTAATTAATCTTGAACCTGCTCTCTCTCTCTCTCTCTCTCT 1410
| | | | |
QY 1367 GCTCAGACCTAAGATCGAGTTTAAAGCTGAAAATGAGGAAAAATTCACATCTATTCACACC 1426
| | | | |
Db 1411 GCTTACGACTCGAATCGAGTTTAAAGCTGAAAATGAGGAAAAATTCACATCTATTCACACC 1470
| | | | |
QY 1427 AGTTGCTACCATCTCACTACCTACCTACCAAGGCTCTTGGTTCCACGGGTACACCAAGT 1486
| | | | |
Db 1471 AGTTGCTACCATCTCACTACCTACCTACCAAGGCTCTTGGTTCCACGGGTACACCAAGT 1530
| | | | |
QY 1487 GGTGATGATCTATGCAAGAGAGCGGCAATGCTGGAACCAATTAAGAGGGCTTGTGTTGG 1546
| | | | |
Db 1531 GGTGATGATCTATGCAAGAGAGCGGCAATGCTGGAACCAATCTTAAGGGCTTGTGTTGG 1590
| | | | |
QY 1547 ATTGGCCCAAGAGATTAACATGATTTCTGAGTACAAAGTGAAGCAATGGGACCGAAGAAATTA 1606
| | | | |

Db 1591 ATTAGCCTCTGAGAACCAATGATCCTGGAGTACAGTGAAGTTCAAGCTTTAAGAGTA 1650

Qy 1607 TATAGTGGGGTAGCCTAGCTGAATG 1632

Db 1651 ATGCTTGGGATACGTAGTGAATG 1676

RESULT 15

```

US-10-424-599-12021/c
; Sequence 12021, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovallik David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5323)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 12021
; LENGTH: 2582
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110862C.1
US-10-424-599-12021

```

Query Match 64.6%; Score 1136.8; DB 18; Length 2582;

QY	2	TCTCTCTTATTCCTTTTGTA--TTTCATTCATCTTAACTCTTGTGTGAAAAATATATGTC	59
Db	2508	TCCTCTTCAATCCCTTTATATACTTTATAGTTTCACTTAACTGTTGGGAAAAATGTC	2449
QY	60	ATCGAAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCCGAGACTGAATTCAGTCC	119
Db	2448	ATCGAAGATTTTAAAGTTGAGAGTCTTAATGTGAAGTACACCGAGACTGAATTCAGTCC	2389
QY	120	GTTGACAACTACGAAACCAACCACTTGTTCACGAAACAGAAATGGCACTTACAGTG	179
Db	2388	GTTGACAACTACGAAACCAACCACTTGTTCACGAAACAGAAATGGCACTTACAGTG	2329
QY	180	ATTGTCAAAACCCAAATCTGTCAAAATCGAATTTAAAAACAATCATGTTCCTAAATTA	239
Db	2328	ATTGTCAAAACCCAAATCCGTCAACTACCAATTTAAAAACAACCAATCTTCAAAATTG	2269
QY	240	GGGGTAAATGCTTGTGGGTTGGGGTGGAAACAACGGCTCAACCTCACCGGTGTGTATT	299
Db	2268	GGGGTAAATGCTTGTGGGTTGGGGTGGAAACAACGGCTCAACCTCACCGGTGTGTATT	2209
QY	300	GCTAACCGAAGGGCATTTTCATGGGCTTCAAAAGACAAAGATTCAACAAGCCAAATTACTTT	359
Db	2208	GCTAACCGAAGGGCATTTTCATGGGCTTCAAAAGACAAAGATTCAACAAGCCAAATTACTTT	2149
QY	360	GGCTCCCTCAACCCAAAGCCTCAGTATCCGAGTTGGGCTTCCAGGGGAGAGAAATCTAT	419
Db	2148	GGCTCCCTCAACCCAAAGCCTCAGTATTCGAGTTGGGCTTCCAGGGGAGAGAAATCTAT	2089
QY	420	GGCCCATTTCAAGACCTGTCTTCAATGGTTAAACCTTGACGACATTTGTGTTGGGGATGG	479
Db	2088	GGCCCATTTCAAGAGATGTGTCTTCAATGGTTAACTTGACGACATTTGTGTTGGGGATGG	2029
QY	480	GATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAAAGTGTTGACATCAT	539
Db	2028	GATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAAAGTGTTGACATCAT	1969
QY	540	TTGAGAAAGCAGTTGAGGCTTATCATGGAATCCATGCTTCACTCCCGGAATCTATGAC	599

Db	1968	TTGGAGAAGCAGTTGAGGCCCTTACATGGAATCCATGTTCCACTCCCGGAATCTAAGAC	1909
Qy	600	CCGGATTTTCATGCTGCGCAACCAAGAGACGTGCCAAACCGTCATCAAGGGACAAG	659
Db	1908	CCGGATTTTCATGCTGCGCAACCAAGAGACGTGCCAAACCGTATTAAGGGACAAG	1849
Qy	660	CAAGAGCAAGTTCAAACAATCATCAAGAGATCAAGGCGTTTAAAGAAAGCCACCAAGTG	719
Db	1848	CAAGAGCAAGTTCAGCAAAATCTATCAAGAGATCAAGGCGTTTAAAGAAAGCCACCAAGTG	1789
Qy	720	GACAAAGTGGTTGTACTGTGGACTGGCAACAAGAGAGTACAGTAAATTTGGTTGGGC	779
Db	1788	GACAAAGTGGTTGTCTGTGGAAGTCCCAACAAGAGGGGTATATGCAATTTGGTTGTAGGC	1729
Qy	780	CTTAATGACACCAATGAGGAATCTTTGGCTGTGGAACAAGAAATGAGGCTGAGATTTCT	839
Db	1728	CTTAATGACACCAATGAGGAATCTTTGGCTGTGGAACAAGAAATGAGGCTGAGATTTCT	1669
Qy	840	CCTTCCACTTGTATGACCATGTCTTGTATATGGAATAATGTCTTTCATTAATGAGAC	899
Db	1668	CCTTCCACTTGTATGATGCTATGTGCTGTGATGGAATAATGTCTTTCATTAATGAGAC	1609
Qy	900	CCTCAGAAACCTTTTGTACAGAGGCTGATTTGATTTTCCATTCGGAGAGAACACTTTGAT	959
Db	1608	CCTCAGAAACCTTTTGTACAGAGGCTGATTTGATTTTCCATTCGGAGAGAACACTTTGAT	1549
Qy	960	GGTGAATATCACTTCAAGATGTGTCAAGACCAACCAATGAAATCTGTGTGTGATTTCTT	1019
Db	1548	GGTGAATATCACTTCAAGATGTGTCAAGACCAACCAATGAAATCTGTGTGTGATTTCTT	1489
Qy	1020	GTGGGGGCTGTATCAAGCCAAATCTATATGTCACTTACCAACATCTGGGAAACAATGAT	1079
Db	1488	GTGGGGGCTGTATCAAGCCAAACATCTATATGTATGTTACCAACATCTGGGAAACAATGAT	1429
Qy	1080	GGTATGATCTTTTCGGCTCCACAACCTTCGCTTCCAAAGGAAATCTCCAAAGCAAGCTT	1139
Db	1428	GGTATGATCTTTTCGGCTCCACAACCTTCGCTTCCAAAGGAAATCTCCAAAGCAAGCTT	1369
Qy	1140	GTTGATGATATGTGCAACAGCAATGCATCTCTATGAGCTGTGTGAACATCCAGACAT	1199
Db	1368	GTTGACGATATGTGCAACAGCAATGCATCTCTATGAGCTGTGTGAACATCCAGACAT	1309
Qy	1200	GTTGTTCTTATTAATATATGTGCTTACTAGTGGGGAACAGCAAGAGGC	1247
Db	1308	GTTGTTCTTATTAATATATGTGCTTACTAGTGGGGAACCAACCATGTGTC	1261

Search completed: June 8, 2005, 11:25:54
Job time : 1101.05 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 21:02:47 ; Search time 300.905 Seconds
(without alignments)
9570.620 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctctatctctcttctgt.....aatgtagtataattctgtc 1760

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	994.2	56.5	1959	4	US-09-727-628-1 Sequence 1, Appl
2	984	55.9	1931	3	US-09-118-442-10 Sequence 10, Appl
3	984	55.9	1931	3	US-09-677-064-10 Sequence 10, Appl
4	384.8	21.9	1578	4	US-09-248-796A-3131 Sequence 3131, Ap
5	349.4	19.9	1231	3	US-09-397-787-34 Sequence 34, Appl
6	340.2	19.3	1602	4	US-09-734-237B-72 Sequence 72, Appl
7	340	19.3	1605	4	US-09-734-237B-74 Sequence 74, Appl
8	160.6	9.1	3546	3	US-09-118-442-15 Sequence 15, Appl
9	160.6	9.1	3546	3	US-09-677-064-15 Sequence 15, Appl
10	159	9.0	3546	3	US-09-118-442-14 Sequence 14, Appl
11	159	9.0	3546	3	US-09-677-064-14 Sequence 14, Appl
12	89.4	5.1	77626	4	US-09-949-016-12608 Sequence 12608, A
13	62.2	3.5	294	4	US-09-913-294A-4684 Sequence 4684, Ap
14	48.4	2.8	1335	4	US-09-902-540-5256 Sequence 5256, Ap
15	48.4	2.8	34316	4	US-09-902-540-1257 Sequence 1257, Ap
16	44.6	2.5	1141	4	US-09-806-708B-22 Sequence 22, Appl
17	37.6	2.1	7218	1	US-08-232-463-14 Sequence 14, Appl
18	37.6	2.1	601	4	US-09-949-016-55935 Sequence 55935, A
19	37.2	2.1	601	4	US-09-949-016-117305 Sequence 117305, A
20	37.2	2.1	601	4	US-09-949-016-117304 Sequence 117304, A
21	37.2	2.1	77388	4	US-09-949-016-13486 Sequence 13486, A
22	37.2	2.1	115388	4	US-09-949-016-14881 Sequence 14881, A
23	37.2	2.1	127771	4	US-09-949-016-14882 Sequence 14882, A
24	36.6	2.1	512	4	US-09-270-767-5181 Sequence 5181, Ap
25	36.6	2.1	512	4	US-09-270-767-20463 Sequence 20463, Ap
26	36.2	2.1	1098	3	US-09-221-017B-928 Sequence 928, App
27	35.6	2.0	251769	4	US-09-949-016-13185 Sequence 13185, A

C	28	35.6	2.0	251769	4	US-09-949-016-13185	Sequence 13186, A
C	29	35.6	2.0	266748	4	US-09-949-016-13187	Sequence 13187, A
C	30	35.6	2.0	266748	4	US-09-949-016-13188	Sequence 13188, A
C	31	35.2	2.0	1584	4	US-09-602-777A-419	Sequence 419, App
C	32	35.2	2.0	2310	5	PCT-US95-072A-4	Sequence 4, Appl
C	33	35.2	2.0	2310	5	PCT-US95-072A-4	Sequence 4, Appl
C	34	35.2	2.0	3387	5	PCT-US95-072A-5	Sequence 5, Appl
C	35	35.2	2.0	3387	5	PCT-US95-072A-5	Sequence 5, Appl
C	36	35.2	2.0	5000	3	US-09-104-070-1	Sequence 1, Appl
C	37	35	2.0	505	4	US-09-621-976-15639	Sequence 15639, A
C	38	35	2.0	2025	4	US-09-107-532A-2547	Sequence 2547, Ap
C	39	34.8	2.0	399	4	US-09-621-976-8976	Sequence 8976, Ap
C	40	34.8	2.0	601	4	US-09-949-016-148492	Sequence 148492, A
C	41	34.8	2.0	2518	3	US-09-433-699-3	Sequence 3, Appl
C	42	34.8	2.0	2527	4	US-09-949-016-4169	Sequence 4169, Ap
C	43	34.8	2.0	13489	4	US-09-949-016-15911	Sequence 15911, A
C	44	34.6	2.0	291	4	US-09-313-294A-518	Sequence 518, App
C	45	34.4	2.0	1446	4	US-09-902-540-5772	Sequence 5772, Ap

ALIGNMENTS

RESULT 1
US-09-727-628-1
Sequence 1, Application US/09727628
Patent No. 6791013
GENERAL INFORMATION:
APPLICANT: Armstrong, Katherine
APPLICANT: Hey, Timothy D
APPLICANT: Folkerts, Otto
APPLICANT: Smith, Kelley A
APPLICANT: Hopkins, Nicole L
TITLE OF INVENTION: MAIZE MIP SYNTHASE PROMOTER
FILE REFERENCE: 50597
CURRENT APPLICATION NUMBER: US/09/727, 628
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: US 60/168, 612
PRIOR FILING DATE: 1999-12-02
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1959
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (137)..(1699)
US-09-727-628-1

Query Match 56.5%; Score 994.2; DB 4; Length 1959;
Best Local Similarity 77.8%; Pred. No. 2.3e+299;
Matches 1200; Conservative 0; Mismatches 343; Indels 0; Gaps 0;

QY	54	AGTTTCATGAGAAATTTTAAAGTGTCTTAATGTAAAGTACACCAAGCTAGATT	113
DB	137	ATGTTTCATGAGAAAGCTTCGCGTGAAGCCCGACGTGCGGTACCGCCGACGAGATC	196
QY	114	CAGTCGCTGTAACCTAGCAACCAACCACTTTGTTCAGCAAGCAAGATGACCTAT	173
DB	197	GAGTCGAGATACCGGTACGACACGACGAGCTGTGTCACGAGCCCAAGCCGCTCC	256
QY	174	CAGTGGATTTGCAAAACCAATCTGTCAATACGAATTTAAACCAACATTCATGTTCT	233
DB	257	CGCTGGGTGTCGCGCCCAAGTCGTCAGTCAACTTCGACGACGACCGCGTCC	316
QY	234	AAATTAGGGTAAATCTGTGCTGGGTGCAAAACAGCTCAACCTCAACCGTGT	293
DB	317	AAGCTCGGGGTATCTGTGGGTGGGAGGCAACAGGGTCCACGCTGACGCTGGG	376
QY	294	GTTATTGCTTAACGAGAGGCAATTCATGCGGTACAAAGCAAGATTCAAGCAAT	353
DB	377	GTCATTGCCACAGGAGGGAGATCTCATGGGCGACCAAGGACAAAGTGCAGCAAGC	436

```
OY 354 TACTTGGCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGCTCTTCCAGGGAAGAGA 413
DB 437 TACTAGGCTCCCTCAACCAAGCTTCCACATCAGAGTGGAGCTACACAGGGAAGAG 496
OY 414 ATCTATGCCCATTCGAAGAGCTGCTTCCATGGTTAAACCTGACGACATTTGTTGGG 473
DB 497 ATATATGCGCGCTTCAAGAGCTCTCTACCATGTGTAAACCCAGACGACCTTGTGTGGA 556
OY 474 GGATGGGATATCAGCAACATGAACTGCTGATGCCATGGCCAGAGGCAAAAGTGTTCAC 533
DB 557 GCGTGGGACATCAGCAGCATGAACTGGGAGATGCCATGACAGGAGCCCAAGGTGTGAGC 616
OY 534 ATCGATTTGTCAGAGAGCTTGAAGGCTTACATGGAATCCATGCTTCCATCCCGGATC 593
DB 617 ATTGACCTGACAGAGAGCTCAGGCGCTTACATGAGATCCATGCTGCACTTCCCGGTGC 676
OY 594 TATGACCCGGAATTCATTTGCTGCCAACGAAGAGAGCGTGCCAAACATCATCAAGGCG 653
DB 677 TATGATCCGGACTTCAATGCGCGCTAACGAGGCTCTCGGCCAAATGTCAATCAAGGCG 736
OY 654 ACCAAGCAAGAGAGAGCTTCAACAAATCATCAAAAGACATCAAGCGCTTTAAGGAAGCCAC 713
DB 737 ACCAAGAAAAGAACAGGTGAGAGAGATCATCAAGAGATATCAGGAGTTTAAAGAGAAGAAC 796
OY 714 AAAAGTGAACAAGGTGTTGTAAGTGTGAGCTGCGCAACAGAGAGGTACAGTATTTGGTT 773
DB 797 AAAAGTGAACAAGGTATTTGCTGTGAGCTGCAACAACTGAAAGGTACAGCATATATGT 856
OY 774 GTGGGCTTAAATGACCAATGAGAAATCTTGTGCTGTGTGAGACAGAAATGAGGTGAG 833
DB 857 GCTGGCTCAACGACACATGAGAAATCTGTGCGCATCTGTGAGACAAAGAGCGGAG 916
OY 834 ATTTCTCTTCCACCTTGATGCCATTTGCTGTGTATGGAATAATTTCTTCAATAT 893
DB 917 ATCTGCGCATCAACATATATGCTATGCTGTGTGACGAGAGGCTGTGCTTCACTCAAT 976
OY 894 GGAAGGCTCAGAACACTTTTGTACAGAGGCTGATGATCTTGTCCATCCGAGAAACAT 953
DB 977 GGAAGGCTCAGAACACTTTTGTGCTGTGCTGTGATTTGATCTTGTATCAAGAACTGCG 1036
OY 954 TTGATTTGTGAGAGTACTTCAAGAGTGTCAAGCAACCAAAATGAATCTGTGTGTTGAT 1013
DB 1037 CTGATCGGTGTGACGACTTCAAGAGTGTGAGCAACCAAGATGAATCGTCTGTGTTGAT 1096
OY 1014 TTCTTTGTGGGGCTGTATCAAGCAACATCTATGTGAGTTTCAACCATCTGGGAAC 1073
DB 1097 TTTCTTTGTGTGTGAGTAAAGCCACCTGATTTGAGCTACAAACCACTTGGGAAC 1156
OY 1074 AATGATGATGATATCTTTCGAGCTCACAACCTTTGCTTCAAGGAATCTCAAGAGC 1133
DB 1157 AACGAGCGGATGAACTGTCTGCTCCCTCAAACTTCAAGTCCAAAGAGATCTTCAAGAGC 1216
OY 1134 AACGTTGTGATGATATGATCAACAGCAATGCCATCTCTATGAGCTGTGAAATCCA 1193
DB 1217 AACGTTGTGATGATGATCAACAGCAATGCCATCTCTATGAGCTGTGAAATCCA 1276
OY 1194 GACCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGGGACAGCAAGAGCATATGAT 1253
DB 1277 GATCATGTTGTTGTTATTAAGTATGTGCTTACGTAGGGGACAGCAAGAGCATATGAT 1336
OY 1254 GAGTACACTTCAAGATATTCATGGGTGGAAGAGACCAATGTTTGTGCAACACATGC 1313
DB 1337 GAGTACACTTCAAGATATTCATGGGTGGAAGAGACCAATGTTGCTGCAACACATGC 1396
OY 1314 GAGATTTCCCTCTTACGCTCTCTATATCTTGTGACTTGTCTTTCGAGCTCAAC 1373
DB 1397 GAGGACTGCTCTCGCGGACCGGACCATCTGATCTGTGTCTCTGTGCTGAGCTCAAC 1456
OY 1374 ACTAAGATGAGTTTAAAGTGAAGGAAAATGAGGAAAATTCATCTCAATCCAGCTTGT 1433
DB 1457 ACCAGGATCAGTTAAACCTGAGGGAAGGAAAGCAAGTTCACATCTTTCACACCGGTGCGC 1516
```

```
OY 1434 ACCATCTCAGCTACCTCACCAGGCTCTGTGTTCCACCGGGTACCAAGTGTGAT 1493
DB 1517 ACCATCTTAGGTACTCACCAGGCAACCATGTGTTCCACCGGCAACCGGTGTGAAC 1576
OY 1494 GCATTGTCAAAGCAGCGTCAATGTGTGAAAACATATAGAGGCTGTGTGTGATTTGCC 1553
DB 1577 GCTCTTGAAAGCAGAGGCGCATGTGTGAGAACATCAAGAGGCTGTGTGTGCTGTGCC 1636
OY 1554 CCAGAGATTAACATGATCTCTGAGTACAAAGTGAAGCATGGAGC 1596
DB 1637 CCAGAGAACACATGATCTCTGAGTACAAAGTGAAGCATGGAGC 1679

RESULT 2
US-09-118-442-10
; Sequence 10, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytrate Metabolism in
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-10

Query Match 55.9%; Score 984; DB 3; Length 1931;
Best Local Similarity 76.1%; Pred. No. 3.5e-296;
Matches 112; Conservative 0; Mismatches 380; Indels 0; Gaps 0;

OY 5 TCTTTATTCCTTTTGTATTTTCATTCATTTTAACTCTTGTGAAAATAATGTTATGCA 64
DB 50 TCGCTACCTCGCTTGCATTCATTCGAAAAGAGGGAAGCAAGATGTTCTATGCA 109
OY 65 GAATTTTAAGGTTAGTGTCTTAATGTGAAGTACACGAGACTGATTCAGTCCGTGTA 124
DB 110 GAGCTTCCGCTCGAGAGCCCTCAGCTGCGGTAGGCGCCGATGAGATCGAGTCCGAGTA 169
OY 125 CAATACGAAACCAACGAACTTGTTCACAGAAACAGAAATGCACTTACAGTGAATTGT 184
DB 170 CCGGTACGACAGACGAGGCTGTACAGAGGCAAGAGCGGCTCACGCTGAGTGGT 229
OY 186 CAATCCCAATGTGTCAATACGAATTTAAACCAACATTCATGTTCTTAAATTAAGGGGT 244
DB 230 CCGCCCAAGTCCGTCGAATCACTTCGGAACCGAAGCCGCGCTCCCAAGCTCGGGGT 289
OY 245 AATGCTGTGGGTGTGGGTGGAACCAAGGCTCAACCTCAACCGGTGGTGTATTGCTTA 304
DB 290 GATGCTGTGGGTGTGGGTGGAACCAAGGCTCAACGCTGAGGCTGTGGGTGATTTGCCAA 349
OY 305 CCGAGAGGCAATTTGATGAGGCTTACAAAGAACAGATTCACAGCAATTAATTTGAGCTC 364
DB 350 CAGGAGGGAATCTATGAGGCGGACCAAGAGCAAGGTGACAGACCACTACTACGAGCTC 409
OY 365 CCTCAACCAAGCTCAGCTATTCGAGTGGGTCTTTCAGAGGAGAGAAATCTATGCCCC 424
DB 410 CTTCAACCAAGCTCAGCTATTCGAGTGGGTCTTTCAGAGGAGAGAAATCTATGCCCC 469
```

425 ATTCAAGAGCGTCTTCAATGGTTAACTGACGACATTTGGTTGGGGATGGATAT 484
470 GTTCAAGAGCGTCTTCAATGGTTAACTGACGACATTTGGTTGGGGATGGATAT 529
485 CAGCAACATGAACCTGGCTGATGTCATGGCCAGAGGCAAGAGTTGGATCGATTGCA 544
530 TGGCAACATGAACCTGGCTGATGTCATGGCCAGAGGCAAGAGTTGGATCGATTGCA 589
545 GAAGCAGTTGAGGCTTCAATGGATTCATGCTTCACTCCCGAATCTATGACCCGGA 604
590 GAAGCAGTTGAGGCTTCAATGGATTCATGCTTCACTCCCGAATCTATGATCCGGA 649
605 TTTCAATGCTGCAACCAAGAGAGGCTCCAAACACGTCATCAAGGCAACAAAGCAAA 664
650 CTTCATCGCGGCTTCAACCAAGAGAGGCTCCCGCAACATGTCATCAAGGCAACAAAG 709
665 GCAAGTTCAACCAATCATCAAGAGAGGCTTCAAGAGAGGCTTCAAGAGAGGCTTCA 724
710 ACAAGTGAAGAGATCATCAAGAGATTCAGAGAGTTTAAAGAGAGAGAGAGAGAGCA 769
725 GGTGTTGTTAGTGTGAGCTGCAACACAGAGAGTACAGTAATTTGGTTGGGCTTAA 784
770 GATAGTTGTTGTTGAGCTGCAACACAGAGAGTACAGTAATTTGGTTGGGCTTCA 829
785 TGACACCATGAGAGATCTTGGCTGCTGTCGACAGAGAGTACAGTAATTTGGTTGG 844
830 CACACAGTGAAGAGATCTTGGCTGCTGTCGACAGAGAGTACAGTAATTTGGTTGG 889
845 CACCTTGAATGAGATCTTGGCTGCTGTCGACAGAGAGTACAGTAATTTGGTTGG 904
890 AACATATATGCTATGCTGCTGTCGACAGAGAGTACAGTAATTTGGTTGG 949
905 GAACACTTTTGAACAGAGGCTGATGATCTTGGCTGCTGTCGACAGAGAGTACAGTA 964
950 GAACACTTTTGAACAGAGGCTGATGATCTTGGCTGCTGTCGACAGAGAGTACAGTA 1009
965 AATGATCTTCAAGAGAGTACAGAGAGTACAGTAATTTGGTTGGTTGGTTGGTTGG 1024
1010 TACAGCTTCAAGAGAGTACAGAGAGTACAGTAATTTGGTTGGTTGGTTGGTTGG 1069
1025 GAGCTGATCAAGAGAGTACAGAGAGTACAGTAATTTGGTTGGTTGGTTGGTTGG 1084
1070 TGGCTGAATTAAGAGAGTACAGAGAGTACAGTAATTTGGTTGGTTGGTTGGTTGG 1129
1085 GAATCTTTGGGCTCAACAACTTTCCGTTCAAGAGAGTACAGTAATTTGGTTGGTT 1144
1130 GAACCTGCTGCGCTCAACAACTTTCCGTTCAAGAGAGTACAGTAATTTGGTTGGTT 1189
1145 TGAATATGTCAGAGAGTACAGAGAGTACAGTAATTTGGTTGGTTGGTTGGTTGGTT 1204
1190 TGAACATGCTGAGAGAGTACAGAGAGTACAGTAATTTGGTTGGTTGGTTGGTTGGTT 1249
1205 TGTATTAATGAATGAGAGTACAGAGAGTACAGTAATTTGGTTGGTTGGTTGGTTGGTT 1264
1250 TGTATCAATGAATGAGAGTACAGAGAGTACAGTAATTTGGTTGGTTGGTTGGTTGGTT 1309
1265 AGAGATATTCATGAGAGTACAGAGAGTACAGTAATTTGGTTGGTTGGTTGGTTGGTT 1324
1310 AGAGATATTCATGAGAGTACAGAGAGTACAGTAATTTGGTTGGTTGGTTGGTTGGTT 1369
1325 CTTAGCTGCTCTATATATTTGGAATTTGCTTCTGCTGAGCTGAGATGAATGCA 1384
1370 CTTCGCGGAGCTATATCTTGTATCTTGTGCTGAGCTGAGATGAATGCA 1429
1385 GTTAAAGCTGAATTAAGAGAGTACAGAGAGTACAGTAATTTGGTTGGTTGGTTGGTT 1444
1430 GCTGAAGAGTGAAGAGAGTACAGAGAGTACAGTAATTTGGTTGGTTGGTTGGTTGGTT 1489
1445 CTACCTCAACAGAGCTCTCTGTTTCAACCGGATACAGAGTGAATGCAATTTGCAAA 1504
1490 TTAATCAACAGAGAGCTCTCTGTTTCCCTGAGACACCGGATGAAGCTCTGAGCAAA 1549

QY 1505 GCAGGCGTGAATCTGGAATAATATGAGGGCTTGTGTTGGATTTGGCCCAAGAGATA 1564
DB 1550 GCAGAGGGGATGCTGGAATAATATGAGGGCTTGTGTTGGATTTGGCCCAAGAGATA 1609
QY 1565 CATGATTTCTGAGTACAGAGTGAAGCATGGGAC 1596
DB 1610 CATGATTTCTGAGTACAGAGTGAAGCATGGGAC 1641

RESULT 3
US-09-677-064-10
; Sequence 10, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Calt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phyate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677, 064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-10

Query Match 55.9%; Score 984; DB 3; Length 1931;
Best Local Similarity 76.1%; Pred. No. 3, 5e-296;
Matches 1212; Conservative 0; Mismatches 380; Indels 0; Gaps 0;

QY 5 TCTTTATCTCTTGTGTAATTTCAATCTTATCTTTGTGTAATAATATTTGATCGA 64
DB 50 TCGTACCTCGCTTGTGTAATTTCAATCTTATCTTTGTGTAATAATATTTGATCGA 109
QY 65 GAATTTTAAGGTTGAGTGTCTTATGTAAGTACAGAGTGAATTCAGTCCGTGA 124
DB 110 GAGCTTCGCGCTGAGAGAGGCTTCAATCTTATCTTTGTGTAATAATATTTGATCGA 169
QY 125 CAATTAAGGTTGAGTGTCTTATGTAAGTACAGAGTGAATTCAGTCCGTGA 184
DB 170 CCGGTACGACAGAGAGGCTGTAACAGAGGCAAGAGCGGCTGCACTGAGGTGCT 229
QY 185 CAATTAAGGTTGAGTGTCTTATGTAAGTACAGAGTGAATTCAGTCCGTGA 244
DB 230 CCGGCAAGTTCGCTCAATCTTATGTAAGTACAGAGTGAATTCAGTCCGTGA 289
QY 245 AATGCTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAG 304
DB 290 GATGCTTGTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTG 349
QY 305 CCGAGAGGCAATTTGATGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAG 364
DB 350 CAGGAGGAGATCTGATGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGG 409
QY 365 CTTACCCCAAGCTTCAATCTTATGTAAGTACAGAGTGAATTCAGTCCGTGA 424
DB 410 CTTACCCCAAGCTTCAATCTTATGTAAGTACAGAGTGAATTCAGTCCGTGA 469
QY 425 ATTCAAGAGCGTCTTCAATGGTTAACTGACGACATTTGGTTGGGGATGGATAT 484
DB 470 GTTCAAGAGCGTCTTCAATGGTTAACTGACGACATTTGGTTGGGGATGGATAT 529

```
QY 485 CAGCAACATGAACTGGCTGATGCAATGGCCAGGGCAAAAGTGTGTCATGATTTGCA 544
    |||||
Db 530 TAGCAACATGAACCTGGCCGACTCATGACAGGGCCAAAGTGTGATATTTGACTCTGA 589
QY 545 GAAGCAATTGAGGCTTACATGGAATCCATGCTTCCACTCCCGGAATCTATGACCCGGA 604
    |||||
Db 590 GAAGCAGCTCAGGCCCTCATGATGAGTCCATGTGTGCACTCCCGGATCTATGATCCGGA 649
QY 605 TTTCATGTGTCACCAAGAGAGGTCGCAACCAACGTCATCAAGGGCCAAACAGCA 664
    |||||
Db 650 CTTCATCGGGCTAACCAAGGCTCTCGGGCCCAAGTGTCTATCAACAGGGCCAAAGAAAGA 709
QY 665 GCAGATTCAACAAATCATCAAAAGACATCAAGCGTTTAAAGAAAGCCACCACCAAGTGA 724
    |||||
Db 710 ACAGGTGAGCAGATCATCAAGATATACGGAGTTTAAAGAGAAAGAAACAAAGTGAACA 769
QY 725 GGTGTTGATCTGTGAGCTGCCAACAAGAGGTACAGTAATTTGGTTGGGCTTTAA 784
    |||||
Db 770 GATATGTTGTGTGTGACTGCAAAACACTGMAAGTATAGCAATGTGTGGCTGTCTCAA 829
QY 785 TGACACCATGAGAAATCTTGGCTGTGTCGACCAAGAAATGAGGCTGAGATTTCTCTTC 844
    |||||
Db 830 CGACACGATGAGAAATCTAAGCTGTGTGACCAAGAAAGAGCGAGATATCAACATC 889
QY 845 CACTTGTATGCTCATTTGTTGTATGGAATGTTCTTTCATTTAATGGAAGCCCTCA 904
    |||||
Db 890 AACACTATATGCACTTGTCTGTGTCATGGAAGGGTCCGCTTCATCAATGGAAGGCCCA 949
QY 905 GAACACTTTTGTACCAAGGCTGATGATCTTCCATCGCGAGAAACATTTGATTTGGT 964
    |||||
Db 950 GAACACCTTTGTGCTGGCTGATGATCTTCTATTAATAAAACAATGCTTATATTTGG 1009
QY 965 AGATACCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGTTGATTTCTTGTGG 1024
    |||||
Db 1010 TGACACTTCAAGAGTGTGACAGCAAGTAAATCTGTGTTGATTTCTTGTGG 1069
QY 1025 GGCTGATCAAGCAACATCTATAGTCATTACCAACATCTGGGAAACATGATGTA 1084
    |||||
Db 1070 TGCTGAATTAAGCCACTCAATGTGTGCTACCAACCTTGGGAAACATGATGTA 1129
QY 1085 GAATCTTTGGCTCCCAAACTTTCCGTTCCAAAGAAATCTCCAAAGAGCAATGTTGA 1144
    |||||
Db 1130 GAACCTGTCTGCCCTCAAAACATTCAGGTCCAAAGAGATCTCCAAAGAGCAATGTTGA 1189
QY 1145 TGATATGTCACAGCAATGCATCTCTATAGAGCTGTGTAATCCAGCAATGTTG 1204
    |||||
Db 1190 TGACATGCTGTGAGCAATGCAATCTCTATAGAGCCGCGAGCAATCCGATCATGTCT 1249
QY 1205 TGTATTAAGTATGTCCTTAAGTAGGGGACAGCAAGAGCCATGATGATGATCACTTC 1264
    |||||
Db 1250 TGTCAATCAAGTATGTGCGGTAGTGGGAGACAGCAAGGGCTATAGAGATACACTC 1309
QY 1265 AGAGATTAATCAAGGTGTGAAAGACCACTTTTGTGCAACAACATCGAGGATTCCT 1324
    |||||
Db 1310 AGAGATCTTCAATGAGGGGCAAGAACCACTGTGTGCAACAACCTGTGAGAGACTCG 1369
QY 1325 CTTAAGTGTCTATTAATTTGGAATTTGTCCTGTGTGAGCTGAGCACTAGATGATGA 1384
    |||||
Db 1370 CCTCGCGCACTATCATCTTGATGTGTGCTCTTGGCTGAGCTAGCAACAGATGATCA 1429
QY 1385 GTTTAAAGCTGAAAAATGAGGAAAAATTCACATCAATTCACCAAGTTGATACCTCTCG 1444
    |||||
Db 1430 GCTGAAAGCTGAGGAGAGAGAAATTCACATCTCTTCCACCGGTGTGCAACATCTTGAG 1489
QY 1445 CTACTCAACCAAGCTCTCTGTGTTCAACCGGATACCAAGTGTGATGATGCTCAAA 1504
    |||||
Db 1490 TTACTTCAACCAAGGCAACCTGTGTTCCCTGTGCAACCGGTGTGAGAGCTGTGCAAA 1549
QY 1505 GGAGGTGCAATGCTGTGAAACATATAAGAGGCTGTGTGATTTGGCCCGCAGAGAAATA 1564
    |||||
Db 1550 GCAGAGGGGATGTGAGAAACATATGAGGCTGTGCTTGGGCTGTGCGCCAGAGAAACA 1609
    |||||
```

```
QY 1565 CATGATTTCTGAGTACCAAGTGAAGCATGGGAC 1596
    |||||
Db 1610 CATGATCTTGAGTACCAAGTGAAGCCTAAGTGGC 1641
    |||||

RESULT 4
US-09-248-796A-3131
; Sequence 3131, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keich Weinlock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT FILING DATE: 1999-02-12
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3131
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-3131

Query Match 21.9%; Score 384.8; DB 4; Length 1578;
Best Local Similarity 56.0%; Pred. 4e-109;
Matches 812; Conservative 0; Mismatches 602; Indels 36; Gaps 3;

QY 163 ATGCGACCTATCAGTGAATGTCAAAACCAATCTGTCAAAATAGCAATTTAAACCAACA 222
    |||||
Db 113 ATGCAATAGTAATATTCATCGTTACACCACTGTAGTACATGATTTAAAGTTGATTT 172
    |||||
QY 223 TCCATGTTCCTAATTTAGGGTATGCTTGTGGTGTGGGTGGAACACGGCTCAACC 282
    |||||
Db 173 TAAAGTCCCTAAGGTTGTTATTTGTAAGTCGATTTGGGAGTAACACGTTACTAAT 232
    |||||
QY 283 TCACCGGTGTGTTATTTCTAAGGAGGCAATTTCAATGGGCTACAAAGACAAGATTC 342
    |||||
Db 233 TGTATGTGTCACATTAAGCAACAAACAAATATTTCTTTGAAACAAAGAGTGTG 292
    |||||
QY 343 AACCAAGCAATTAATTTGAGCTCCCTACCAAGCTCAGCT-----ATCCGAGTTGGT 396
    |||||
Db 293 TTAACCAATTAATGATGTTCTGTACTCAAGCATTCACATGTCATAATCGGTGATA 352
    |||||
QY 397 CTTTCAAGGAGGAATTAATGATGCTTCAAGAGCTGCTTCAATGTTAACTCTG 456
    |||||
Db 353 AAGAAATGAGGATGATTTACGTTCCATTTAATCTCATTTGTTCCATGTTAACTCA 412
    |||||
QY 457 ACGACATTTGTTGGGGATGGGATATCAGCAACATGAACTGAGTCATGAGCCA 516
    |||||
Db 413 ATGATTTGTTGATGATGTTGGATTAATTAATGTTGTTTCCATTTAGCAACCAATGAAGA 472
    |||||
QY 517 GGGCAAGATGTTGACATTCATTTGCAAGAGCAAGTTGAGGCTTACATGGAATTCATGC 576
    |||||
Db 473 GAGCTAAAGTCTTGATGATTTACTTTACAAAGCAAGTTGATTCATTTGAGAAACAGA 532
    |||||
QY 577 TTCACTCCCGGAATCTATAGCCCGATTTCAATTTGCTGCAACCAAGAGAGGTGCA 636
    |||||
Db 533 AACCTTTGAATCAATTAACCTGATTTCAATTTGCTTGAACCAAGTAAAGTGA 592
    |||||
QY 637 ACAAGTCATCAAG-----GGCAAAAGCAAGCAAGTTT 672
    |||||
Db 593 ACAATGTTTAAACCAAGTCAATGATGATTAATACTGATTAATTAATGAGCCGAGCTTG 652
    |||||
QY 673 AACCAATCATCAAGACATCAAGGCTTTAAGAAAGCCACCAAGTGAACAAGTGTG 732
    |||||
Db 653 AAAAATGAGAAAGATTCAGAGATTTCAAGGCCAAAGAAAGAAATTAATTAAGTTATTA 712
    |||||
QY 733 TACTGTGACTGCAACACAGAGATACAGTAATTTGCTTGTGGCTTATATGACCA 792
    |||||
```


Db 713 TTTTATGACGTCTAACTAGAAAGATATGCTGATGTTTCCAAATGTCATGACACTG 772
Qy 793 TGGAGATCTCTTGCGCTGTGAGACAGAAATGAGCTGATGTTCTCTCCACTGTG 852
Db 773 CTGACACTTGTATCAAGAGATATTAAGAAATCAACGAAAGAAATGCTCCAAATGCTCT 832
Qy 853 ATGCGATGCTGTGTTATGAAATGTTCTTCAATTAATGGAAGCCCTCAGAACACTT 912
Db 833 TTGCGCTGTGTTATCTTGAAGAAATGTCATATATATATGTTTACCCAAACATCAT 892
Qy 913 TTGTACCAAGGCTGATGATCTTCCATCGCAGAAACACTTTGATTTGTGAGATGACT 972
Db 893 TTGTTCCCGGTATATGATGATGCTGAAATATACGACTCATTCATGTTGGTGTGATGAT 952
Qy 973 TCAAGATGCTGACCAAAATGAATCTGTGTGTTGATTTCTTGTGGGGCTGTGTA 1032
Db 953 TCAAGTACGCTCAACAAATTAATTAATCACTGTGATGCTCAATTCCTGTCGATGCTGTA 1012
Qy 1033 TCAAGCCAAATCTATGATGATGATTAACACCATCTGGGAAACATGATGATGATCTTT 1092
Db 1013 TCAAACTCTTCTATATGCTTCTTATATATCACTGGGTAAACATGACGTTACATTTAT 1072
Qy 1093 CGGCTCCAAACTTTCCGTTCCAGGAAATCTCAAGAGCAACGTTGTTGATGATATGG 1152
Db 1073 CATCACCAAAACATTTAGATCTAAGAAATTTCCAAACATCTGTTGTTGATGATATTA 1132
Qy 1153 TCAACAGCAATGCAAT-----CCTCTATAGCCTGTGTAACATCCAGACATGTTGTTG 1206
Db 1133 TTGAATCCAAAGAAATTAATGTAACAAAGAAATCTGTGCAAAAGTTGATCACTGATATCG 1192
Qy 1207 TTATTAATGATGTCCTTACGTAGGGGACAGCAAGAGACATGATGATGATCACTTTCAG 1266
Db 1193 TCATTAATATCTTGTCCAGCTGTGTTGATTTCTAAAGTTGCAATGATGATGATTTTCA 1252
Qy 1267 AATATATTCATGCTGTGAAAGAGACCAATGTTTTCACAAACATGATGATGATTTCCCTCT 1326
Db 1253 AATTAATGTTGGGTGTGTCACAAACAAATTAATGATTCACATGTTTGTGAAGATTTCTGC 1312
Qy 1327 TACGCTGCTCTATATCTTGTGATGTTGCTCTTCTTGTGAGCTCAGACTAGAAATGAGAT 1386
Db 1313 TTGCTACACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1372
Qy 1387 TTAAGCTGAAATATGAGGAAATTCACATTCACCCAGGATGATGATGATGATGATGATG 1446
Db 1373 TCAAGGCTCTGTTAATCTGATTAATGATGATGATGATGATGATGATGATGATGATGAT 1432
Qy 1447 ACCTCAACAAAGCTCTCTGTTTCCACCGGTAACACAGTGTGATGATGATGATGATGATG 1506
Db 1433 ACTGCTCAAGGCTCATTAAGCAAGACAGGATTCAAACCTATCAACGATTTAAACAAAC 1492
Qy 1507 AAGCTGCAATGCTGGAAGAAATTAATGAGGCTTGTGTTGATGATGATGATGATGATGAT 1566
Db 1493 AAGCTCAACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1552
Qy 1567 TGAATCTGCA 1576
Db 1553 TAAGATTTGA 1562

RESULT 5
US-09-397-787-34
; Sequence 34, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16

; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 1231
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-34
Query Match 19.9%; Score 349.4; DB 3; Length 1231;
Best Local Similarity 62.0%; Pred. No. 3.8e-98;
Matches 587; Conservative 0; Mismatches 356; Indels 4; Gaps 2;
Qy 637 ACAAGCTCATCAAGGCGCAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 696
Db 31 ACAAGCTCATCAAGGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 90
Qy 697 CGTTTAAGAGAGCCAAAGTGAACAGAGTGTGTTAGTGTGAGCTGTCACACAGAGA 756
Db 91 ACTTCGGGTCTAGCGCGGGCTGAGCAAAAGTATATGCTGTGAGACGGGAAACAGCGAGC 150
Qy 757 GGTACGATATTTGTTGGGCTTTAATGACCAATGAGAAATCTTGTGCTGTG 816
Db 151 GCTTCTGTAGAGTGTTCAGGCTTCACAGACAGCCGAGAACTGTGCGCACCATTTG 210
Qy 817 ACAAGAAATGAGCTGATTTGCTTCCACTGTGATGCAATGCTGTGTTATGAGAA 876
Db 211 AGCTCGGTG--GAGGTGTGCTTCCAGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCG 267
Qy 877 ATGTTCTTTCATTAATGAGAGCCCTCAGAAACATTTTACAGAGGCTGATGATCTTG 936
Db 268 GCGTGTCTTCTCATATGGGTCTCCGACAGAACCTGTGTCGCGAGCTTTGAGCTCG 327
Qy 937 CCATGCGGAGAAACATTTGATGTTGTGAGATGATCTTCAAGAGTGTGACCAAAATGA 996
Db 328 CGTGACAGACCGGCTTTTGTGGGCGAGATGACTTCAAGTCAGCCAGACCAAGTCA 387
Qy 997 AATCTGTGTTGTTGATTTCTTGTGGGGCTGTATCAAGCAACATGATGATGATGAT 1056
Db 388 AGTCTGTGTTGATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
Qy 1057 ACAACCATCTGAGAAACATATGATGATGAA--TCTTTCGCTTCCAAACTTTCGCTTCC 1115
Db 448 ACAACCATCTGAGAAACATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 507
Qy 1116 AAGAAATCTCCAAAGACCACTGTTGATGATGATGATGATGATGATGATGATGATGAT 1175
Db 508 AAGAGGTGTCCAAAGGCAACGTGTGAGACGATGTCAGACGACCAAGCTCTAT 567
Qy 1176 GAGCTGTGATCATCCAGCATGTTGTTTAAATGATGCTTAAAGGAGGAGC 1235
Db 568 ACGCCCGGAGAGGCTGACCATGCTGTGATCAATGATGATGATGATGATGATGATGATG 627
Qy 1236 ACGAAGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1295
Db 628 ACGAAGCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
Qy 1296 GTTTTGCACACATGCGAGGATTCCTCTTAAGCTGCTCTTATTAATCTTGAACCTTGTG 1355
Db 688 GTGCTGACCAACAGTGTGAGAGCTGCTGTGCGCGCACCATGATGCTGAGACCTAGCG 747
Qy 1356 CTTCCTGCTGAGCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1415
Db 748 CTGCTGACCGAGCTGTGACGCGCTGAGCTTTCGACCTGACATGAGACCCGAGCGCAG 807
Qy 1416 TCATTCACACCGAGTGTCTACATCTCTCAGCTACCTCAGCAAGGCTCTTGTGTTCCAGCG 1475
Db 808 ACCTTTCACCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 867
Qy 1476 GGTACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1535
Db 868 GGCACGCGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 927
Qy 1536 GCTTGTGTTGATTTGGCCCGAGAAATTAACATGATTTCTGATGATCA 1582

Db 928 GCGTCGCGGGGCTCCCGCACAGAACCATGCTCTGGAAACAGAA 974
RESULT 6
US-09-734-237B-72
; Sequence 72, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734, 237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494, 921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
; LENGTH: 1602
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-734-237B-72

Query Match 19.3%; Score 340.2; DB 4; Length 1602;
Best Local Similarity 55.4%; Pred. No. 3.4e-95;
Matches 775; Conservative 0; Mismatches 573; Indels 51; Gaps 4;

QY 224 AAATTAGGGGTAATGCTTGTGGGTGGGAGAAACACGCTCAACCTTCACCGGTGT 293
DB 133 AAACATAGGAATATGCTCATTTGGGTTAGGTGGCAACATAGCTCCACTTTAGTGGCTTG 252
QY 294 GTTATTCCTAACCGAGAGGCAATTCATAGGCTTACAAAGCAATTAAGCCAT 353
DB 253 GTATTGGCGAATAGACAAATGTGAGATTCAAACTTAAGAGAGCGGTTAAGCAACCAAC 312
QY 354 TACTTGGCTCCCTCAACCCAGCTCAGCTATCCAGTGGGTCTCTCC--AGGAGAG 410
DB 313 TACTTCGGCTCATGACTCAATGTTCTAAGCTTGAAGCTGGATTCGATCGGAGGGGAAAT 372
QY 411 GAAATCTATGCCCATTCAGAGCTGCTTCAATGTTAACCTTCAGCAATGTGTTT 470
DB 373 GAGCTTATGCTCTTTTAATCTCTGTTGCCATGCTTAAAGCCAAACGACTTTGTGCTC 432
QY 471 GGGGGAATGGGATTCAGCAACATGAACCTGGCTGATGCCATGCCAGGCAAAAGGTGT 530
DB 433 TCTGTGGGACATCAATACGCAATCTTAACGAGCTATGCAAGAAATCAAGTTCTC 492
QY 531 GACATCGATTTTCAGAGAGCTTGAAGCTTACATGGAATTCATGCTCCCGGA 590
DB 493 GAATATGATCTGCAACAGCTTTGAAGGGAAGATGTCTTGTGAAGCTCTTCTCTTC 552
QY 591 ATCTATGACCCGGAATTCATGCTGCAACCAAGAGAGCGTTCACCAACGTCATCA- 649
DB 553 ATTACTACCTCGATTTCAATGAGCTTAATCAAGATGAGAGCCAAATACGATCAAT 612
QY 650 -----GGGCACAAGCAAGCAAGTTCACCAATATCATC 683
DB 613 TTGGATGAAAAAGCAACGTAAACAGAGGGGTAACTGAGCCATCTGAGCAACATA 672
QY 684 AAAGCATCAAGCGCTTAAAGAGAGCCACCAAGTGAAGAGTGTGTAAGTGTGAGT 743
DB 673 CCGGATATCCAGATTTCAAAAGAGAAAGCCCTTGATTAAGTAAATCGTTTGGACT 732
QY 744 GCGCAACAGAGAGTACAGTAATTTGTTGGGCTTAATGACACATGAGAAATCTC 803
DB 733 GCAAAATACGAGAGGTAGTAAGATATCTCCGTGTTAATGACCAATGAGAAACCTC 792
QY 804 TTGGCTGCTGTGACGAATAGGCTGATTTCTCTTCACTTGTATGCTATGCT 863
DB 793 TTGCACTTATTAAGATGACCATGAAAGATTTGCTCTTCAAGATCTTTTGCAGCAGCA 852

QY 864 TGTGTTATGAAAAATGTTCTTTCAATTAATGAAGCCCTCAGAACACTTTTGTACAGGG 923
DB 853 TCTATCTTGAAGGAGGTCCCTATATTAATGTTTCAACGAGAAATACCTTTGTTCCCGGC 912
QY 924 CTGATTTATCTTGCATGCGAGAGAACCTTGAATGTTGGAGATGATCTTAAGAGGTGT 983
DB 913 TTGGTTACGCTGGCGAGCATGAGGGTACATTCATTTGGGGAGAGATCTCAAGCGGGA 972
QY 984 CAGACCAAAATGAATCTGTGTGTTGATTCTTGTGGGGGCTGTATCAAGCCAACA 1043
DB 973 CAACCAAGTTGAATCTGTCTTGGCCAGTTCTTAAGTATGACAGATATTAACCGGTTC 1032
QY 1044 TCTATATGCTAGTTACAAACATCTGGAAAACAATGATGATGAATCTTTGGCTCCACA 1103
DB 1033 TCCATTTGATCCTTATTAACATTTAGGCAATATGACGGTTAACTTATCTGCTCAAAA 1092
QY 1104 ACTTTCGCTTCAAGGAATCTCCAAAGCAACGTTGTGATGATGATGATCAACAGCAT 1163
DB 1093 CAATTTAGGTCTTAAGAGATTTCCAAAGTCTGTCAATGATGATCATATGCGCTTAAT 1152
QY 1164 GCCATCTCT-----ATGAGCTGTGAACATCCAGACCAATGTTGTTATTAAGTAT 1217
DB 1153 GATATCTTGTACATATGATTAATGAGTAAAGTTGAACCACTGATTTGTATCAAAAT 1212
QY 1218 GTGCTTACGTAGGGGACAGAGAGCCATGATGATGATCACTTCAAGATATTCATG 1277
DB 1213 ATGAAGCCCGTGGGAGCTCAAAAGTGGCAATGACAGATATTAACATGATGATGAT 1272
QY 1278 GGTGAAAGAGACCATTTGTTTGCACAAACATGCGAGATTTCCCTTGTAGCTGCTCT 1337
DB 1273 GGTGGCCATTAACCGGATTTCAATTCACATGTTTGGCAAGATCTTTTACTGGCTACG 1332
QY 1338 ATTATCTTGAAGCTTGTCTTCTGAGCTCAGCACTAGATGATGATTTAA----- 1390
DB 1333 TTGATCATGATCTTTTATGATGATGATGATTTTATCAAGATGCTCTTATTAAGAGT 1392
QY 1391 -----AGCTAAATGAGGAAATTCATCTATTCACCCAGTTGCTACATCTCTC 1442
DB 1393 GACCCAGTTAAAGAGATGCTGGCAAAATTCGAAATCTTTATCCAGTTTAACTTCTTG 1452
QY 1443 AGCTACCTACCAAGGCTCTGTGTTCCACCGGTACACAGTGGTAAATGATTTGCA 1502
DB 1453 AGTTACTGGTTAAAGCTTCATTAACAAAGACCAAGATTTACCCGTTAAATGGCTTAA 1512
QY 1503 AAGCAGCGTCAATGCTGAAAAACATATGAGGCTTGTGTGATGAGCCCAAGAGAT 1562
DB 1513 AAGCAAAAGCCGCTTGAAGAAATTTTAAATGATTTGATTTGATTTGCTCTCAAAA 1572
QY 1563 AACATGATTTCTGAGTACA 1581
DB 1573 GAATTAAGATTCGAAGAGA 1591

RESULT 7
US-09-734-237B-74
; Sequence 74, Application US/09734237B
; Patent No. 6818752
; GENERAL INFORMATION:
; APPLICANT: Rozzell, J. David
; APPLICANT: Bui, Peter
; APPLICANT: Hua, Ling
; TITLE OF INVENTION: SYNTHETIC GENES FOR ENHANCED EXPRESSION
; FILE REFERENCE: B583:40608
; CURRENT APPLICATION NUMBER: US/09/734, 237B
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/494, 921
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74
; LENGTH: 1605
; TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic gene derived from *Saccharomyces cerevisiae* myo-inositol
 ; OTHER INFORMATION: -1-phosphatase synthase, having numerous codons replaced with other
 ; OTHER INFORMATION: a encoding the same amino acids to reduce free energy of folding,
 ; OTHER INFORMATION: and a gly codon inserted after the initiating met codon
 US-09-734-23B-74

Query Match 19.3%; Score 340; DB 4; Length 1605;
 Best Local Similarity 55.5%; Pred. No. 3,9e-95;

Matches 773; Conservative 0; Mismatches 570; Indels 51; Gaps 4;

234 AAATTAGGGGTAATGCTTGGGTTGGGTTGAAAGCAAGGCTCAACCCCTACCGGTTGGT 293
 196 AAGCTGGGATATCTGCTGATCGGCTGGGTTGTAACAACGCTCTACTCTGTTGATCT 255
 294 GTTATTTCTAACCGAGGAGGCAATTTGAGGCTTACAAAGAACAAAGATTCAACAGCCAAAT 353
 256 GTTCTGGCAACAAACAAACGTAAGATTCCAGACTTAAAGAGGTTTAAACAGCCGAAC 315
 354 TACTTTGGCTCCCTCAACCCCAAGCTCATCTCATGTTGGTCTTCC--AGGAGAG 410
 316 TACTTTGGTCTTATGACTGAGTCTTACTCTGAGCTGGGCAATTGATGCTGAAGGTAAC 375
 411 GAAATCTATGCCCATTTCAAGAGCCGTTCCAAATGTTTAAACCTGACGACATTTGTT 470
 376 GACGTTTACGCTTCCTTCACTCTCTGCTGCGATGATCTCCGAACGACTTCGTTGT 435
 471 GGGGGATGGGATATCAGCAACATGAACCTGCTGATGTCATGAGCCAGGCAAGGTTT 530
 436 TCTGGTGGGATATCAACAACCGGATCTGTAGAAAGCAATGAGGTTCTCAGTTCTG 495
 531 GACATGATTTGGCAGAGCAAGTTGAGGCTTACATGAAATTCATGCTTCACTCCCGGA 590
 496 GAAATGATCTGCAACAGGCTGGAAGGCTAAGATCTCTGGTTAAGCACTGCCGTC 555
 591 ATCTATGACCGGATTTCAATGCTGCCCAACGAAGAGGCTCAACAAACGTCATCAAG 650
 556 ATCTATACCCGGATTTTATCGAGCTTACAGAGCAAGAGTGTAAACAATGTTATCAAC 615
 651 GGCACAAACGAGCAAGCAAGT-----CAACAAATCATC 683
 616 CTGAGCAAGAAAGGTAAGCTTACTACCGGTGAATGAGCTCACTGACGATACCT 675
 684 AAAGACATCAAGCGTTTAAAGAAAGCCACCAAGTGAACAGGTTGTTACTGTGACT 743
 676 CCGTATATCAGAACTTCAAGAGGAAAGCGCACTGGAACAAAGTTATCTACTGTGACT 735
 744 GCAACACAGAGAGGTACAGTAATTGGTTGGGCTTAAATGACCAATGAGAAATTC 803
 736 GCTTACACTGAACGTTACGTAGAGTATCCCGGGTGAACGATCTATGAAAACTG 795
 804 TTGGCTGCTGGGACAGAAATGAGCTGAGATTTCTCTTCCACTTGTATGCTTGTCT 863
 796 CTGCAATCTATCAAGAACACACAGAGAAATCGCTCCGTCACATCTTCTGCTGTGA 855
 864 TGTGTTTGAAGAAATGCTCTTCAATTAATGAAGCCCTCAAGCACTTTTGTACAGGG 923
 856 TGTATCTCTGAAGGCTTACGTAATCAACGCGCTCCGCAAGAACTTTGTAACCGGGT 915
 924 CTGATGATCTTGGCCATGCGAGGAACACTTTGATTTGGTGGAGATACCTTCAAGAGTGT 983
 916 CTGGTACACTGGCTGAACACGAAGGTACCTTCATGCTGGTGAACATCTGAATCTGGC 975
 984 CAGACCAAAATGAATCTGTGTTGTTGATTTCTTGTGGGGCTGTATCAAGCCACA 1043
 976 CAGACTAAATGAATCTGTACTGAGCAAGTTCTCTGTTGACGCTGTATCAACCGGTT 1035
 1044 TCTATATGATGTTACCAACATCTGGGAAACAAATGATGATGATTTTTCGCTCCACA 1103
 1036 TCTATGCTTCTTATTAACCACTGGGTTAAACAAGGCTTACCAACCTGTCTGTCGAA 1095
 1104 ACTTTCCTTCCAGGAATTCACAGACACGTTGTTGATGATGATGTTCAACGCAAT 1163

DB 1096 CAGTTCGCTTAAAGAAATCTCTAATCTCTGTAATGAGACATCATCGTTCTAAC 1155
 1164 GCCATCT-----CTATGAGCTGTGTGAACATCCAGACCATGTTGTTTATTAAGTAT 1217
 1156 GACATCTCTAACAACCAACATCGGGTAAAGATGATCACTGTATCGTTATCAAAATAC 1215
 1218 GTTCCTTACGTAGGGGACAGCAAGAGGCAATGATGATACACTTCAAGATATTCATG 1277
 1216 ATGAAACCGGTTGGTATTTCTAAAGTTGCTATGAGAGAAATCTACTCTAAGCTGATG 1275
 1278 GGTGAAAGAGCAACCTGTTTGTGCAACACATGCGAGATTTCCCTTTAGCTGCTCT 1337
 1276 GCGGTCACACACCGATCTCTTATCCAAAGCTTTGGAAGACTCTTCTGCTGCTACCCG 1335
 1338 ATTATCTTGAGTCTGCTCTTCTGCTGAGCTCAG-----CACTAGAATC 1382
 1336 CTGATCATCAGCTGTGTGTTATGACTGAATTTCTGACCCGTTATCTTAAAGAAAGTT 1395
 1383 GAGTTTAAAGCTGAATATGAGGAAATTCATCTATTCACCCAGTTCATCCATCTTC 1442
 1396 GACCGGTTAAAGAAATGCTGGCAAAATTCGAAACTTCTACCCGTTCTGACCTTCTG 1455
 1443 AGTACCTCACCAGGCTCCTCTGCTTCCACCGGGTACACAGTGGTGAATGATGCTCA 1502
 1456 TCTTACTGCTGAAGAGCTCCGCTGACTGCTCAAGCTTCCACCGGTTAACGCTGAAC 1515
 1503 AAGCAGCGTCAATGCTGGAAACATTAATGAGGCTTGTGTTGATTTGGCCCCAGAAAT 1562
 1516 AAACAGCGTACCGCTGTGAAACCTTCCGCTGTGCTGATCGGCTGCGTCCGACGAAC 1575
 1563 AACATGATTTCTGA 1576
 1576 GAACCTGCGTTTGA 1589

RESULT 8

US-09-118-442-15

; Sequence 15, Application US/09118442B
 ; Patent No. 6197561

GENERAL INFORMATION:

; APPLICANT: Martino-Catt, Susan J.

; APPLICANT: Wang, Hongyu

; APPLICANT: Beach, Larry R.

; APPLICANT: Wang, Xun

; APPLICANT: Bowen, Benjamin A.

; TITLE OF INVENTION: Genes Controlling Phytochrome Metabolism in

; TITLE OF INVENTION: Plants and Uses Thereof

; FILE REFERENCE: 0706

; CURRENT APPLICATION NUMBER: US/09/118,442B

; EARLIER FILING DATE: 1998-07-17

; EARLIER APPLICATION NUMBER: 60/055,446

; EARLIER FILING DATE: 1997-08-11

; EARLIER APPLICATION NUMBER: 60/055,526

; EARLIER FILING DATE: 1997-08-08

; EARLIER APPLICATION NUMBER: 60/053,944

; EARLIER FILING DATE: 1997-07-28

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 15

; LENGTH: 3546

; TYPE: DNA

; ORGANISM: Zea mays

US-09-118-442-15

Query Match 9.1%; Score 160.6; DB 3; Length 3546;

Best Local Similarity 64.5%; Pred. No. 7e-39; Mismatches 84; Indels 84; Gaps 1;

Matches 305; Conservative 0; Mismatches 84; Indels 84; Gaps 1;

DB 1032 ATCAAGCCAAACATCTTATGCTAGTTACACATCTGGGAAACAAATGATGATGATCTT 1091
 2716 ATGAGGCCCACTCAATGCTGATGATCAACCACTTGGGAAACAAAGATGATGATGATCTG 2775

Qy	1092	TCGGCTCCAAACATCTTCOCGTTCCAGGAATCTCCAAAGACACGGTTGTGATGATG	1151
Db	2776	TTGCGCTTCCAAACATTATAGGTCACAGGAGATCTCCAAAGACACGGTGTGATGACATG	2835
Qy	1152	GTCAACAGCAATGCGCATCTCTATGAGCGTGTGAACATCCAGACCAATGTGTGTTATT	1211
Db	2896	GTTTCGAGCAATGCGCATCTCTATGAGCGCGGCGAGCATCCGATCAATGTGTGTATC	2895
Qy	1212	A-----	1212
Db	2896	AAGGTCTGTAGCTGATCTTTCACCTCGTTAAAGTTGACATATGCAAGCAGATTTACA	2955
Qy	1213	-----AGTATGAGCTTACGTAAGGGGACAGCAAGAGACC	1247
Db	2956	TTGAAACTTTCACACTCTTTTGTGGCAGTATGTGCCATCGTGGAGACAGCAAGGGCT	3015
Qy	1248	ATGATGATGATACACTTCAGAGATATTCATGSGTGGAAAGACACCATTGTTTGCACAAAC	1307
Db	3016	ATGGACGAGTACACCTCGAGATCTTCATGSGGCGGCAAGAACCAACATGTGCTGCACAAAC	3075
Qy	1308	ACATGCGAGATTCCTCTTATGCTGCTCTTATTTATCTTGAACCTTGCTCTTCTGCTGAG	1367
Db	3076	ACCTGTGAGGACTCGCTCCTCGCGGCACCTATCATCTTGTGATCTGAGTCTCTTGGCTGAG	3135
Qy	1368	CTCAGCACTAATATGAGACTTTAAAGCTGAATAATAGGGGAAATTCACATCAATT	1420
Db	3136	CTCAGCACTAGATTCAGCTTAAAGCTGAGGAGAGTGAAGACCCCAAGT	3188

```

RESULT 9
US-09-677-064-15
; Sequence 15, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; TITLE OF INVENTION: Genes Controlling Phylate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-15

```

	Query Match	Similarity	9.1%	Score 160.6	DB 3	Length 3546
	Best Local	Similarity	64.5%	Pred. No. 7e-39		
	Matches	305	Conservative	0	Mismatches 84	Indels 84
					Gaps 1	
Qy	1032	ATCAAGCCAACTATATAGTCAGTTACCAACCATCTGGGAAACAATGATGATGAATCTT				1091
Dd	2716	ATGCAAGCCCACTCATATGTCAGTCACCACTTGGGAAACAGATGCAATCACTG				2775
Qy	1092	TCGGCTCCACAACTTTCGTTCCAAAGAAATCTTCAAGACCAAGTGTGATGATATG				1151
Dd	2776	TCTGCGCTTCAAAACATTCAGGTCCAAAGAGATCTTCAAGAGCAAGTGTGATGACATG				2835
Qy	1152	GTCACACGCAATGCATCTCTATATAGCTGTGTAACTCCAGACCATGTTGTTTATT				1211
Dd	2836	GTTCTGACCAATATCCATCTCTATATAGCTGTGTAACTCCAGATCATGTCGTTGTATC				2895

Oy	1212	A-----	1212
Db	2896	AAGCTCTGTTAGCTGATCTTTCACCTCGTTAAAAAGTTGACATATGCAAGGCAGATTTTACA	2955
Oy	1213	-----AGTATGTGCTTTACGTAGGGGAGCAGCAGAGAGACC	1247
Db	2956	TTGAAACTTGTCACTCTTTTGTGTGACATATAGTGCCGTACCTGGGAGACACAGAGGGCT	3015
Oy	1248	ATGATATAGTACACTTACAGAGATTATCATGGTGGAAAGACACATGTGTTTTCACAAC	1307
Db	3016	ATGACGACGATACCTCTAGAGATTTTATATGGCGGCAAGAACACATCTGTGCTGCACAC	3075
Oy	1308	ACATGCGAGGATTCCTCTTAGCTGCTCTATATCTTGAATTGGACTTGTCCTTCTTGCTGAG	1367
Db	3076	ACCTGTAGGAGCTGCTCTCTCGCGCACATATCACTTATCTGTGGCTCTTGGGCTGAG	3135
Oy	1368	CTCAGCACTGAATCGAGTTTAAAGCGAAATAATAGGAAAAATTCACCTATT	1420
Db	3136	CTTCGACACAGAACTCCAGCTGAAGAGCTGAGGAGAGGTTAAAGAGCCCTCCCAAGT	3188

```

; RESULT 10
US-09-118-442-14
; Sequence 14, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu C.
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytoate Metabolism in
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 3546
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-14
```

Query Match	9.0%: Score 159; DB 3; Length 3546;
Best Local Similarity	79.7%: Pred. N.2e-38;
Matches 200; Conservative	0; Mismatches 50; Indels 1; Gaps 1;
Qy 446	GATTAACTGACGACATTTGTGTTGGGGGATGAGATATACGACATGAACCTGGGTGA 505
Db 1559	GGTAAACCCAGACGACATTTGTGTTGGAGGGGTGGGACATTAGCAACATGAACCTGGCGGA 1618
Qy 506	TGCCATGGCCACGGGCAAAGGTTTGAATCGATTTCAGAGACAGTTGAGGCTTACAT 565
Db 1619	CTCCATGACCCAGGGCCAAAGGTGTCTGATTAATGACCTCAGAGACAGCTCAGGGCCCTTACAT 1678
Qy 566	GGATTCGATGCTTTCAC-TCCCGGAGTCTATGACCCGGATTCATTCATGCTGCCAACGAG 624
Db 1679	GGAGTTCATGGTGCCACTTCCCGGATCTATGATTCGGAAGCTTATGCGGGCTTAACAGG 1738
Qy 625	AGGAGCGTGCCAAACAGTCATCTCAAGGGGCAACAAAGCAGACAGTTCAACAAATCATCA 684
Db 1739	GCTCTCGGGCCAAACAGTGTCTATCAAGGGGCAACAAAGAAACAGTGGAGGATCATCA 1798
Qy 685	AAGACATCAAG 695

RESULT 14
US-09-902-540-5256
; Sequence 5256, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5256
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-5256

Query Match 2.8%; Score 48.4; DB 4; Length 1335;
Best Local Similarity 43.4%; Pred. No. 0.00039;
Matches 434; Conservative 0; Mismatches 551; Indels 15; Gaps 4;

```
QY 427 TCAAGAGCCTGCTTCCATGTTAAGCCTGACGACATTTGTTGGGGATGAGATATCA 486
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 194 TCAAGAGCTGTGCTCTTGGCGGAGCTTAAGAGCTTCCCTTGGCGCTGGACATCA 253
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 487 GCACATGAACCTGTGATGCGATGCGCAGGCGCAAGGTGTTGACATGATTTGCAGA 546
DB 254 TCCGAGAGACCATACAGAAAGTGGCGGTGCGGTGCTGACGCGACAGCAACTG 313
QY 547 AGCAGTTGAGGCTTACATGAGATCATCTTCCATCCCCGGAATCTATGACCGGATT 606
DB 314 AGAGGTGAGGCGTCTCCAGAGCATCAAGCGGCGGTGCGACGACCTGAGT 373
QY 607 TCATTGCTCCCAACAGAGAGCGTGCACAAAGTCATCAAGGGCACAAGCAAGAAC 666
DB 374 TCGTGGCGCGCATCGAGCGCAACACATCAAGGCCACCAAGACGAC-----CGGAGA 427
QY 667 AAGTTCAACAATTCATCAAGAATCATCAAGCGCTTTAAGAGACCAACAAGTGACAAAG 726
DB 428 GCATGGAAGCGGTGGCGCCAGACATCGGACCTTCAAGAGAGGCTCAAGCGAGCGCG 487
QY 727 TGGTTGATGTGAGTCCCAACAGAGAGTACATGTAATTTGTTGGGCTTATG 786
DB 488 CCGTATGTGTGTGTCAGACAGGTGAGACCTTCGTCGCGTCCGAGTCTTCA--- 544
QY 787 ACACCATGAGAAATCTTGGGCTGTGAGACAGAAATGAGGCTGATTTCTCTTCA 846
DB 545 AAGACGTGGCGCTTTCAGAAAGCGCTGAGAGAACGCCCGGACATCAACCCACCG 604
QY 847 CCTTGATGCAATGCTTGTGTATGAAATGTTCTTTCAATTAATGAGACCTTCAGA 906
DB 605 CCGTGAACCTTACAGGCGGCATCAAGAGAGGCGGTGCTTTCGGAACGCGACGCCAAG 664
QY 907 ACACTTTTGA--CGAGGCTGATTTGATCTTCCATCGGAGAACACTTTGATTTG 963
DB 665 CCAGCGCTGACACGCGCGGCTTCAGAGAGATGCGCAAGAGAGTGGCGCTGCGCG 724
QY 964 GAGATGACTTCAAGATGTGTCAGACCAAAATGAAATCTGTGTTGTTGATTTCTTGTG 1023
DB 725 GCGCGCACTTCAAGAGCGCGCCAGACATGATGAAAGCGGTCACTCGCGCGGCTTAAG 784
QY 1024 GGGCTGTATCAAGCCACATCTATAGTCAATTAACCATCTGGGAAACAAATGATGTA 1083
DB 785 CCGCATGTGTGGGTGTGATGCGTGTCTTCCACCAAACTCTGGGCAACCGGAGAGGG 844
QY 1084 TGAATCTTTGGCTCAAACTTTCTGTTCCAGAGAAATCTTCAAGAGCAAGTGTG 1143
DB 845 AGGTGTGAGCAACCCGCGGCTTCAAGGCGCAAGAGTCAACAAGTGTGAGGTGTG 904
```

```
QY 1144 ATGATATGTCACAGCAATGCAATCCTCTATGAGCTGTGAAATTCAGACATGTTG 1203
DB 905 AACACATCTCTGAGGCC---GACCTGTACCGGAGCTGTACAGAAATGACGGGCAAGG 961
QY 1204 TTGTTATTAAGTATGCTCTTAAGTAGGAGCAGCAGAGAGCCATGATAGTCACTT 1263
DB 962 TGGCCATCCTACTACATCCCGCGCGCGCGAGCGCAAGAGGAGGTTGACAAATCGACA 1021
QY 1264 CAGAAATTTCAATGGTGGAAAGAGACCATTTGTTTCACAAACATGCGAGATTC 1323
DB 1022 TCACCGATGGCTCGGCTATCCGATGACATCAAGTCAACTTCTCTGCGGACTTCA 1081
QY 1324 TCTTATGCTCTCTATTAATCTTGAAGTGTCTTCTTGTGAGCTCAGACTAGAAATG 1383
DB 1082 TCTGCGCGCGCGCGTGTGTCGACATCGGCTGTCTCGGACTGCGCAAGCGGTG 1141
QY 1384 AGTTAAAGTGAATAATGAGGAAATTCACACTATTCA 1423
DB 1142 AGTGGCGGCGCATCCAGAGATGATGTCTTCTACTTCA 1181
```

RESULT 15
US-09-902-540-1257/c
; Sequence 1257, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1257
; LENGTH: 34316
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1257

Query Match 2.8%; Score 48.4; DB 4; Length 34316;
Best Local Similarity 43.4%; Pred. No. 0.0034;
Matches 434; Conservative 0; Mismatches 551; Indels 15; Gaps 4;

```
QY 427 TCAAGAGCCTGCTTCCATGTTAAGCCTGACGACATTTGTTGGGAGATGATATCA 486
DB 19705 TCAAGAGCTGTGCTCTGCGGAGCTGAAGAGCTTGCGCTTGGCGCTGGACATCA 19646
QY 487 GCAACATGAACCTGGCTGATGCGATGCGCAGGCGAAAGTGTGATCATGATTTGCAGA 546
DB 19645 TCCGTGAGAGCATACAGAAATGAGCGGTGCTCGGGGTGTCAAGCAAGCACTGG 19586
QY 547 AGCAGTGAAGCTTACATGAGATTCATGCTTCACTCCCGGAAATCTATGACCGGATT 606
DB 19585 AGAGGTGAAGCGCTTCTTCCAGAGATCAAGCCCAAGAGCGGTGACAGACCTGAGT 19526
QY 607 TCATTGCTGCAACCAAGAGAGCGTGCACAAAGCTCATCAAGGGCCCAAGCAAGAGC 666
DB 19525 TCGTGGCGCGCATTCAGAGCCAAACATCAAGGCGCACCAAGACGAC-----CGCGAGA 19472
QY 667 AAGTTCAACAATTCATCAAGACATCAAGGCTTTAAGAGGCCCAAAATGAGCAAG 726
DB 19471 GCATGAAGCGGTGCGCAAGCATCCGCACTTCAAGAGAGGCTCAACGAGCGCG 19412
QY 727 TGGTTGATCTGAGACTGCAACAGAGAGGTACAGTAATTTGTTGGGCTTAAATG 786
DB 19411 CCGTATGTGTGTGTCAGACAGCGGTGAGACCTTCCGTCGCGGACGATCTTCA--- 19355
QY 787 ACACCATGAGAAATCTTGTGCTGTGACAGAAATGAGGTGAGATTTCTCTTCA 846
```

```
Db 19354 AGACGCTGGCCGCTTGAGAGAGCGCTGACAGAAACAGCCCGGACATCAACCCACCG 19295
QY 847 CCTTGATGCGCATGCTTGTTGTAATGAGAAATGTCCTTTCATTAATGAGAGCCCTCAGA 906
Db 19294 CGCTGTACACCTACGGCGCCATCAGAGAGGGCGCTGCTTCGGAAACGCCACGCCCAAG 19235
QY 907 ACACCTTTTGA---CGAGGGCTGATGATCTTGCCATCGGAGAGAACCTTTGATGGTG 963
Db 19234 CACGCGTGAACAGCGCGCGCTTCAGAGAGTGGCCAGAGAGTGGCGCTGCGCG 19175
QY 964 GAGATGACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTGATTTCTTGTGG 1023
Db 19174 GCGCGACCTCAAGAGAGGGCAGACGATGATGAGAGCGTCATCGCGCCCGCTCAGG 19115
QY 1024 GGGCTGTATCAAGCCAACTTATAGTCAATTCAACCATCTGGAGAAACATGATGTA 1083
Db 19114 CCGGCATGCTGGGTGTGATGCTGTCTCCACCAACATCTGGGCAACCGGACGGCG 19055
QY 1084 TGAATCTTTCGGCTCCAAACCTTCGTTCCAGAGAAATCTGCAAGAGCAAGTTGTG 1143
Db 19054 AGGTGCTGACAGACCCCGCGCTTCAAGGCCAAGAGATCACCAAGTCGAGGCTGTG 18995
QY 1144 ATGATATGCTCAACAGCAATGCCATCTCTATGAGCTGGTGAACATCCAGACCATGTG 1203
Db 18994 ACACCATCTGCGAGCCC---GACCTGTACCCGAGCTGTACAGAAAGTACGCGCACAGG 18938
QY 1204 TTGTTATTAAGTATGTGCTTACGTAGGGGACAGCAAGAGGCCATGATGAGTACATT 1263
Db 18937 TGGCCATTCACATCTACCGGCCCGCGGCGAGCGAAGGAGGTGGACAAACATCGACA 18878
QY 1264 CAGAGATTTTCATGGGTGAGAAAGACCATTTGTTGCAACACATGCGAGGATTTCC 1323
Db 18877 TCACCGGATGGCTCGGCTATCCGATGACAGATCAAGGTCAACTCTGCGCGACTCCA 18818
QY 1324 TCTTAGCTGCTCTTATTTATCTTGACTTGCTGCTTCTGCTGAGCTCAGACTAGAAATCG 1383
Db 18817 TCCTGGCCGCGCGCTGGTGTGACATCGCGCTGTTCCTGAGCTGGCCAGCGGCTGG 18758
QY 1384 AGTTAAAGCTGAAAAAGAGGAAAAATTCACATCATTTCA 1423
Db 18757 AGTGGCGGCGCATCCAGAGTGGATGCTTCTTACTTCAA 18718
```

Search completed: June 8, 2005, 03:03:24
Job time : 312.905 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 22:54:36 ; Search time 5996.19 Seconds
(without alignments)
11172.621 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctcttattccttctgt.....aatgtaataattctgt 1760

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991.4	56.3	1740	9	CL961092 OaIFCC005
2	819.6	46.6	1828	3	CNS004AMK BX827819 Arabidops
3	799.8	45.4	813	7	CK768601 Gm-r1030-
4	784.4	44.6	2360	1	AJ583520 AJ583520
5	706.8	40.2	988	7	CK277950 EST724028
6	690.8	39.2	782	4	AW348136 AW348136
7	680.8	38.7	754	4	BI968101 Gm830004A
8	658.4	37.4	686	2	AW348857 Gm210010A
9	656.4	37.3	936	7	CK279064 EST725142
10	653.4	37.1	914	7	CK275352 EST721430
11	651.8	37.0	672	2	AW508269 s151h10.Y
12	647.4	36.8	867	6	CB973407 CAB30002
13	642.6	36.5	865	7	CO085839 GR_Ba03E
14	639.4	36.3	898	7	CO113870 GR_Eb014
15	636.2	36.1	906	7	CK320178 L2P08a02
16	633.2	36.0	891	7	CF212628 CGF100065
17	631.8	35.9	860	7	CO072025 GR_Ba30P
18	631.6	35.9	842	7	CF211061 CAB20007
19	628.8	35.7	836	7	CF518404 CAP0007_I
20	608.4	34.6	648	1	AI794750 sb68f07.Y
21	603.4	34.3	860	7	CO082095 GR_Ba46I
22	602.8	34.2	734	6	CA801137 sau01d09.Y
23	601.2	34.2	622	2	AW397496 sg79e12.Y
24	599.2	34.0	670	4	BG044525 Baa29e07.Y

25	597.6	34.0	816	6	CB975215 CB975215
26	595.4	33.8	769	6	CA900545 PCSC1911B
27	594	33.8	978	4	BG838516 GC01_01h1
28	592	33.6	619	2	BE331363 BE331363
29	591.6	33.6	762	2	BE660321 BE660321
30	591.6	33.6	834	7	CK933094 CGF100434
31	591.2	33.6	960	7	CK274458 EST720535
32	590.8	33.6	746	6	CA918410 EST642557
33	590.6	33.6	831	7	CO132349 GR_Eb45C
34	586.8	33.3	830	7	CO107630 GR_Eb003
35	586.2	33.3	909	6	CB345158 CA48EN000
36	586	33.3	841	7	CO075472 GR_Ba36C
37	584.6	33.2	618	2	AW279066 sg07d02.Y
38	580.6	33.0	707	2	BF005128 EST433626
39	579.6	32.9	663	6	CA900542 PCSC17211
40	579.6	32.9	883	6	CD438045 EL01N0508
41	576.2	32.7	781	6	CB343821 CA32EN000
42	573.2	32.6	891	7	CV290142 acf01-10m
43	569.2	32.3	897	7	CF514596 CABud0005
44	567.4	32.2	594	4	BM523521 Baa85a05.Y
45	566	32.2	575	2	AW397453 sg79p01.Y

ALIGNMENTS

RESULT 1	CL961092	1740 bp	DNA	linear	GSS 21-SEP-2004
LOCUS	OaIFCC005463	Oryza sativa Express Library	Oryza sativa	(indica	
DEFINITION	cultivar-group) genomic, genomic survey sequence.				
ACCESSION	CL961092	GI:52376905			
VERSION	CL961092.1				
KEYWORDS	GSS.				
SOURCE	Oryza sativa (indica cultivar-group)				
ORGANISM	Oryza sativa (indica cultivar-group)				
REFERENCE	1 (bases 1 to 1740)				
AUTHORS	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G. K. S., Deng, X. W. and Wang, J.				
TITLE	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis				
JOURNAL	Unpublished (2004)				
COMMENT	Contact: Chen Chen Department of Bioinformatic Beijing Institute of Genomics Chinese Academy of Sciences, Beijing 101300, China Tel: 86-10-80481559 Fax: 86-10-80488676 Email: chenchen@genomics.org.cn Rice genomic sequence. Class: exon-trapped.				
FEATURES	Location/Qualifiers				
source	1..1740 /organism="Oryza sativa (indica cultivar-group)" /mol_type="genomic DNA" /db_xref="taxon:39946" /clone_lib="Oryza sativa Express Library" /note="Oryza sativa exon trapped genomic sequences "				
ORIGIN					
Query Match	56.3% Score 991.4; DB 9; Length 1740;				
Best Local Similarity	78.0%; Pred. No. 1.5e-276;				
Matches 1193; Conservative	0; Mismatches 336; Indels 0; Gaps 0;				
Oy	54 ATGTTTCATCAGAGATTGAAGTTAGTGTCTTAATGTGAAGTACACCGAGCATGATT 113				
Db	1 ATGTTTCATCAGAGATTTCAGGCGGTGAGAGCCGCGCATCGGCGCGGAGATC 60				
Oy	114 CAGTCCGCTGTACACTAGAAACACCGAATCTTGTTCACGAGAACAGAAATGGACACTT 173				

Db 61 GAGTGGAGTACAGACAGACAGAGGCTGTGTGACAGAGACAGACGAGCGGCTCC 120
Qy 174 CAGTGGATTTGCAAAACCCAAATCTGTCAATTCGAATTTAAACCAATCATTTTCT 233
Db 121 CGCTGGGTGTGTCGGCCAGAGTCGCTGCTAATCTTCGGACCAACCAACCTGTCCC 180
Qy 234 AAATAGAGGGTAAATCTTGTGGGTGTGGGGTGAACAACAGCTCAACCTCAACGGTGT 293
Db 181 AAGCTCGGGGTGATGCTGTGGGGGTGGGGGGGCAACAGCGCTCAACGCTGACGGCTGG 240
Qy 294 GTTATTTGCTAACCGAGAGGCAATTCATGGGCTAACAGAGCAAGATTCACAGCCAT 353
Db 241 GTCATGCGCAACAGAGGAGGAATCTCATGGCGCAACAGAGACAGAGTGACAGAACCAAC 300
Qy 354 TACTTTGGCTCCCTCAACCCAGGCTCAGATATCCAGATTGGGTCTTTCAGAGGAAGAGA 413
Db 301 TACTATGTGCTCACTACCCAGGCGTCCACATCAGGGTGGGAGCTTCAACCGGGAGGAG 360
Qy 414 ATCTATGCCCCATTCAGAGGCTGCTCCATATGTTAAACCTGACGACATTTGTTGGG 473
Db 361 ATCTAGCGGCTTCAAGAGCTCTCTGCGCATGTGAAACCTGATGACCTTGTGTCCGG 420
Qy 474 GGATGGGATATGAGCAATGAACTGTGCTGATGCTGATGCTGAGGCGCAAGGTTTGAAC 533
Db 421 GGCTGGGACATTAAGCAATGAACCTGGCTGATGCTATGACCAAGGACAAAGGTACTTGAC 480
Qy 534 ATCGATTTGCAAGACAGATGAGGCTTACATGAAATCCATCTTCCACCGGGAATC 593
Db 481 ATGTATCTGCAAGACAGCTCAGACCTTACATGAGATCCATGTCTCTCCCGGCAATC 540
Qy 594 TATGACCCGAGTTTATTTGCTCTCCAAACCAAGAGAGCGTGCACAAACGTCATCAAGGAC 653
Db 541 TATGACCCGAGTTTATTTGCTCTCCAAACCAAGAGATCCCGGCGAACAATGTCTATCAAGGAC 600
Qy 654 ACAAGCAAGACAGCAATTCACAAATCATCAAGATCAAGAGCGTTTAAAGAACCAAC 713
Db 601 ACCAAGAGAGACAGATGAGAGATCATCAAGACATCAGAGGATTCAGAGAAAGAAC 660
Qy 714 AAGTGGAGCAAGTGTGTCTACTGTGAGACCTGCAACAGAGATGAGTAAATTTGGT 773
Db 661 AAGTGGAGCAAGT 720
Qy 774 GTGGGCTTAAATGACACCATGAGAGATCTCTTGCTGTGTGTGTGTGTGTGTGTGTGT 833
Db 721 GTTGGGCTCAATGACATGAGAGATCTCTTGCTGTGTGTGTGTGTGTGTGTGTGTGT 780
Qy 834 ATTTCTCTTCCACCTTGTATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 893
Db 781 ATATCACCATCGACACTGTATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy 894 GGAAGCCCTCAAGACATTTTGTACAGAGGCTGATGATCTTGTCCATCGCGAGAACACT 953
Db 841 GGAAGTCTCAAGACATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Qy 954 TTGATTTGGAGATGACTTCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1013
Db 901 CTGATTTGGT 960
Qy 1014 TTCTTTGGGGGCTGT 1073
Db 961 TTCTTTGGGGGCTGT 1020
Qy 1074 AATGATGT 1133
Db 1021 AATGATGT 1080
Qy 1134 AAGT 1193
Db 1081 AATGATGT 1140
Qy 1194 GACCATGTTGT 1253

Db 1141 GATCATGTTGT 1200
Qy 1254 GAGTACACTTCAGAGATTTATTCATGGGTGGAAGAGACACATTTGTTTTCACAAACATGC 1313
Db 1201 GAGTACACTTCAGAGATTTATTCATGGGTGGAAGAGACACATTTGTTTTCACAAACATGC 1260
Qy 1314 GAGATTCCTCTTGT 1373
Db 1261 GAGACTCACTTCTTGT 1320
Qy 1374 ACTAGATTCAGATTTAAAGT 1433
Db 1321 ACCAGATTCAGATTTAAAGT 1380
Qy 1434 ACCATCTCAGCTACCTCAGCAAGGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1493
Db 1441 GCCCTGGCAAGAGAGGCAATGTTGAGAAATCATGAGAGGCTGTGTGTGTGTGTGTGTGT 1500
Qy 1554 CCAGAGATTAATCATGATTTCTGAGTACAA 1582
Db 1501 CCCGAGAACCAATGATCTGTGAGTACAA 1529

RESULT 2
CNS04AMK
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA complete sequence from clone GSLTPG1242E06 of Hormone Treated Callus of strain col-0 of Arabidopsis thaliana (thale cress).
BX827819
ACCESSION
BX827819.1 GI:42462602
VERSION
HTC; GSUT_cDNA.
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1828)
REFERENCE
Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat M. Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL
Unpublished
REFERENCE
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV IRRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
source
Location/Qualifiers
1..1828
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"

/db_xref="taxon:3702"
 /clone="GSLRPGH24ZE06"
 /issue_type="Hormone Treated Callus"
 /plasmid="pCMVSPORT 6"
 complement(1..1828)
 /gene="Atg39800"
 ORIGIN

Query Match 46.6%; Score 819.6; DB 3; Length 1828;
 Best Local Similarity 71.2%; Pred. No. 1.6e-226;
 Matches 1129; Conservative 0; Mismatches 444; Indels 13; Gaps 3;

```

31 ATTCTTAATCTTTGTGAAAAAATAATGTTATCGAGAAATTTTAAGTTGATGTCCTAATG 90
35 AATTTTCTTTTCTTCTTCAAAAATTTTATGAGCTTCAAGTTGAGAGCCGAAAG 94
91 TGAAGTACACGAGACTGAGATTCAGTCCGTGTACAACTACGAAACACCGAATTGTTG 150
95 TGAAGTACACAGAGATGAGATTCAGTCTGTGTACGATTACGAGACGAGAGTTGTTG 154
151 AGCAGAA---CAGGAATGGCACTATCAGTGTGTTGTCAAACTCAATCTGTCAATAG 207
155 ATGAGAAAGACCGTCAATGGCACTTACCAATGATGTGAAGCCAAAGACTGTCAATAG 214
208 AATTTAAACCAACATCCATGTTCTTAATTTAGGGGTAAAGCTGTGGGTGGGTTGAA 267
215 ATTTCAAAACCGATATTCGTTGTCCTCAAAATGAGGGTTAGCTGTGGGTGGGTTG 274
268 ACAGCGCTCAACCTCACCGGTGTGTATTGCTAACCGAGAGGCAATTTCAAGGCTA 327
275 ACAATGATCAACTCTCACCGGTGTGTATTGCTAACCAAGAGAAATCTGTGGGCA 334
328 CAAAGACAGAGTTCAACAGCAATTAATCTTTGCTCCCTCACCCAGGCTCACTATCC 387
335 CCAAGACAGAGTTCAACAGCAATTAATCTTTGCTCCCTCACCCAGGCTCACTATCC 394
388 GAGTTGGGTCTTCCAGGAGAGAAATCTATGCCCCCATTTCAAGAGCTGTTCCAAATG 447
395 GTGTGCAATCTTTTACGGGTGAAGATCTATGCTCTTCAAGATCTCTTCCAAATG 454
448 TTAACCTGACGACATTTGTTGGGGGATGGGATATCAGCAACATGAACCTGCTGATG 507
455 TGAATCCGATGATGTTGTTGGAGATGGGATATTAAGTACATGAATTAAGCAGATG 514
508 CCATGGCCAGGCAAGGTTGTTGATGATGATTGCAAGCAAGTTGAGGCTTACATG 567
515 CCATGGCTAGAGCCAGGGTTCTTGTATATGACTTGCAGAAACAGCTCAGGCTTACATG 574
568 AATCATGCTTCACTCCCGGAAATCTATGACCCGATTTCAATGCTGCCAACAAAG 627
575 AAGACATTTGTGCACTCCCTGGGATCTTGTATCTGATTTCAATGCTGTTAAACAGG 634
628 AGCGTCCAAACAGCTCATCAAGGCAAGCAAGCAAGCAAGTTCAACAAATCATCAAG 687
635 CACGTGCTAACACAGTGAATCAAGATCAAGAGAGAAACAGTCAACATCATCAAG 694
668 ACATCAAGCGCTTAAAGAGCCACCAAGATGACAGAGGTGTTGACTGTGAGCTGCA 747
695 ACATCAAGCGCTTAAAGAGAGAAATTAAGTGAATTAAGTGTGTTGTTCTGGAAGGCTA 754
748 ACACAGAGGTTACGATATTTGTTGGGCTTAATGACACCAATGAGATTTCTTGG 807
755 ACACAGAGGTTATGCAATGTGTGTTGGATGAAACATCAAGTGAAGATTTCTATG 814
808 CTGCTGTGACAGAAATGAGGCTGAGATTTCTCTTCCACTTGTATGCTGCTTGG 867
815 ACTGTGTGATGAGGATGAGGCTGAGATTTCTCTTCAACACTTATGAGATTTGCTG 874
868 TTATGAGAAATGTTCTTCTTCAATTAATGAGAGCCCTCAGAACTTTTGTATCAAGGCTGA 927
875 TTCTTGAAGGAAATCCCGTTCAATTAATGAGAGCCCTCAGAACTTTTGTATCAAGGCTTA 934
928 TTGATCTTGTCCATCGGAGAAACATTTGATTTGTTGAGATGACTTCAAGAGTGTGAGA 987
  
```

Db 935 TTGATATGCGCATCAGAAACAAATGTTTATGATGCGTGGAGATATTTCAAGATGGCAGAA 994
 Oy 988 CCAAAATGAATCTGTGTTGTTGATTTCTTGTGGGGGCTGTATCAAGCAATCTTA 1047
 Db 995 CCAAGATGAATCTGTGTTGTTGATTTCTTGTGGGCTGTATCAAGCAATCTTA 1054
 Oy 1048 TAGTCAATTAACCAATCTGGGAAACAAATGATGATTAATTTTCCGCTCCAAACTT 1107
 Db 1055 TTGTAGCTACATCACTTGGGAAACAAAGATGAGATTTCTCAAGTCCACAGACTT 1114
 Oy 1108 TCCGTTCCAAAGAAATCTTCCAGAGCAAGCTTGTGATGATATGTCACACAGCAATGCCA 1167
 Db 1115 TGAATCTAAGAGATTTCTCAAAAGCAATGTGTGAGCATATGTTGCTAGCAATGTA 1174
 Oy 1168 TCCTCTATGAGGCTGTGTAACATCCAGACATGTTGTTATTAATATGTCCTTACG 1227
 Db 1175 TCCTCTTGAAGCTGTGGGAAACACTCGAGC-CCATGTTGTTGTCATACAGAGAAACAAG 1233
 Oy 1228 TAGGGACAGCAAGAGCCATGATGATGATACATTTCAAGATATTTCAAGGTTGAAAG 1287
 Db 1234 TCGACCAATCATCAAGACATGAGGAGTTTAAAGAGAAATTAAGTGAATTAAGTTG 1293
 Oy 1288 GCACCATTTGTTTGCACAAACATGCGAGATTTCCCTCTTACGTGCTCTATTTATCTTGG 1347
 Db 1294 TGCTTCTCGAGCGCTTAACAGAGCTTTATGCAATGTGTCTGTGGATGAAACGATA 1353
 Oy 1348 ACTTGGTCTTCTTGTGACGCTCAGACATGATGATGATTTAA-----GCTGAAA 1398
 Db 1354 CAATGAGAAATCTCATGAGGTGTGTGATAGGATGAGGCTGATCTCTCTTCAACAC 1413
 Oy 1399 ATGAGGAAATTTCAATCATTTCCACCAGTTGCTAATCTCAAGTCTCAACAG 1458
 Db 1414 TTATGAGAAATTTCACTTTTCACTTGTGCGCAACATCTCAGTCACTCACCAG 1473
 Oy 1459 CTCTCTGCTTCCACCGGCTACACAGTGTGATGATGATGATGATGATGATGATGATG 1518
 Db 1474 CACCTCTTGTACCGCGGAACTCCGATGATTAAGAACTGTGAGACAGCGGCTATGC 1533
 Oy 1519 TGAAGAAATTAATGAGGCTGTGTTGATTTGGCCCCAGAGAAATTAATGATTTGAGT 1578
 Db 1534 TGAAGAAATTAATGAGGCTGTGTTGATTTGGCTCCGAGAAACATGATGATGAT 1593
 Oy 1579 ACAATGAAAGATGGGACCGAAGAT 1604
 Db 1594 TCAAGTGAACATGAAGAGCATGAT 1619

RESULT 3
 CK768601 813 bp mRNA linear EST 20-FEB-2004
 LOCUS Gm-r1030-621 Gm-r1030 Glycine max cDNA clone Gm-r1030-621 5', mRNA
 DEFINITION
 ACCESSION CK768601
 VERSION CK768601.1 GI:42722702
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 REFERENCE 1 (bases 1 to 813)
 AUTHORS Periapuram,C.C., Li,L., Wurtele,E.S., Westgate,M.E. and Nikolau,B.J.
 TITLE Molecular determination of soybean composition
 JOURNAL Unpublished (2004)
 CONTACT: Basil J. Nikolau and Cyril C. Periapuram
 Department of Biochemistry, Biophysics and Molecular Biology
 Iowa State University
 2210 Molecular Biology Building, Ames, IA 50011, USA
 Tel: 515 294 9423
 Fax: 515 294 0453

Email: dimas@iastate.edu
Individual base call and confidence value were assigned using the
Phred software (http://www.phrap.org/). Overall sequence quality
assessment and vector trimming were conducted using the Lucy
software (http://www.tigr.org/software/).
This clone was originally generated by the Public Soybean EST
Project (http://129.186.26.94/soybeanest.html)/Shoemaker, R
(rcshoe@iastate.edu).
This clone is available through: Biogenetic Services, 801 32nd Ave,
Brookings, SD 57006. For further information call 605-697-8500 or
contact info@biogeneticservices.com
Seq primer: T7-1 (5' AAT ACG ACT CAC TAT AG 3').
Location/Qualifiers

FEATURES
source

1. 813
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="Gm-r1030-621"
/lab_host="DH10B"
/clone_lib="Gm-r1030"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from
immature cotyledons of greenhouse grown plants
(individual seed fresh weight of 100-300mg). The library
was prepared using the Life Technologies pSuperScript cDNA
library construction kit. Complementary DNA was
synthesized from mRNA using a poly(dT) sequence with a
NotI restriction site. SalI linker adapters were ligated
to the blunt-ended cDNA fragments followed by NotI
digestion. The cDNA fragments were directionally cloned
into the NotI-SalI restriction site of the pSPORT1
vector. The ligated cDNA fragments were transformed into
E. coli Electromax DH10B host cells. This library was
constructed by Dr. Ilya Vodikin and Dr. Anu Khanna. Note
that Gm-r1030 is a re-rack of Gm-cl007."

ORIGIN

Query Match 45.4%; Score 799.8; DB 7; Length 813;
Best Local Similarity 99.8%; Pred. No. 7.1e-221;
Matches 801; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 485 CAGCAACATGAACCTGGCTGATGCCCAAGGCGGCAAGGCTTTGACATGATTTGCA 544
DB 11 CGGCAACATGAACCTGGCTGATGCCCAAGGCGGCAAGGCTTTGACATGATTTGCA 70
QY 545 GAAGCAGTTGAGGCTTACATGAATCCATGCTCCACCTCCCGAATCATGACCCGGA 604
DB 71 GAAGCAGTTGAGGCTTACATGAATCCATGCTCCACCTCCCGAATCATGACCCGGA 130
QY 605 TTTCATTTGCTGCCCAACAGAGAGGCTGCCAACAACATCATCAAGGCGCAACAAGCAAGA 664
DB 131 TTTCATTTGCTGCCCAACAGAGAGGCTGCCAACAACATCATCAAGGCGCAACAAGCAAGA 190
QY 665 GCAAGTTCAACAATATCATCAAGAGCATCAAGGCGTTTAAAGAACCCACCAAGTGGACAA 724
DB 191 GCAAGTTCAACAATATCATCAAGAGCATCAAGGCGTTTAAAGAACCCACCAAGTGGACAA 250
QY 725 GGTGGTTGACTGTGGAGCTGCCAACAAGAGGTAAGTAATTTGGTTGGGGCTTAA 784
DB 251 GGTGGTTGACTGTGGAGCTGCCAACAAGAGGTAAGTAATTTGGTTGGGGCTTAA 310
QY 785 TGACACCATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAGATTTCTCTTC 844
DB 311 TGACACCATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGCTGAGATTTCTCTTC 370
QY 845 CACCTTTGATGCAATGCTTGTGTTATGGAATATGTTCTTTCATTAATGAAAGCCCTCA 904
DB 371 CACCTTTGATGCAATGCTTGTGTTATGGAATATGTTCTTTCATTAATGAAAGCCCTCA 430
QY 905 GAACACTTTTGTACAGAGGCTGATTTGATTTGGCAATCGGAGAAACATTTATATGGTGG 964
DB 431 GAACACTTTTGTACAGAGGCTGATTTGATTTGGCAATCGGAGAAACATTTATATGGTGG 490
```

```
QY 965 AGATGACTTCAAGAGTGTGACAGACCAAAATGAATCTGTGTGGTGAATTTCTTTGGG 1024
DB 491 AGATGACTTCAAGAGTGTGACAGACCAAAATGAATCTGTGTGGTGAATTTCTTTGGG 550
QY 1025 GGCTGGTATCAAGCAATATATAGTACATCAACATCGGGAACATGATGAT 1084
DB 551 GGCTGGTATCAAGCAATATATAGTACATCAACATCGGGAACATGATGAT 610
QY 1085 GAATCTTGGGCTCCACAACTTTCCGTTCCAGAGAAATCTCAAGCAAGTGTGTA 1144
DB 611 GAATCTTGGGCTCCACAACTTTCCGTTCCAGAGAAATCTCAAGCAAGTGTGTA 670
QY 1145 TGATATGTCACAAGCAATGCAATCTCTATGAGCTGTGTAACATCCAGACATGTTGT 1204
DB 671 TGATATGTCACAAGCAATGCAATCTCTATGAGCTGTGTAACATCCAGACATGTTGT 730
QY 1205 TGTATTATAGTATGCTTACATGAGGGAACAGCAAGACCATGATGATGATCACTTC 1264
DB 731 TGTATTATAGTATGCTTACATGAGGGAACAGCAAGACCATGATGATGATCACTTC 790
QY 1265 AGAGATATTCAAGGCTGGAAGA 1287
DB 791 AGAGATATTCAAGGCTGGAAGA 813
```

RESULT 4
AJ583520/c 2360 bp mRNA linear EST 08-JAN-2004
LOCUS AJ583520 PpProcl Physcomitrella patens cDNA clone pp001015066
DEFINITION similar to putative myoinositol-1-phosphate synthase, mRNA
sequence.

ACCESSION AJ583520.1 GI:40781590
VERSION AJ583520
KEYWORDS EST.

SOURCE Physcomitrella patens
ORGANISM Physcomitrella patens

REFERENCE 1 (bases 1 to 2360)
Hone, A., Egener, T., Luecht, J.M., Holterf, H., Reinhard, C., Schween, G.
and Reek, R.
An improved and highly standardised transformation procedure allows
efficient production of single and multiple targeted gene-knockouts
in a moss, Physcomitrella patens
Curr. Genet. 44 (6), 339-347 (2004)

JOURNAL Contact: Schween G
Plant Biotechnology
University of Freiburg
Sonnenstrasse 5, D-70104 Freiburg, Germany.

FEATURES
source
1. 2360
/organism="Physcomitrella patens"
/mol_type="mRNA"
/db_xref="taxon:3218"
/clone="pp001015066"
/tissue_type="protonema"
/clone_lib="PpProcl1"

ORIGIN

Query Match 44.6%; Score 784.4; DB 1; Length 2360;
Best Local Similarity 70.4%; Pred. No. 3.1e-216;
Matches 1111; Conservative 0; Mismatches 451; Indels 16; Gaps 4;

```
QY 23 TTTCATTTCTTAATCTTTGTGAAATATATGTTCAATGAGATTTTAAGTTGAGTG 82
DB 2181 TTTCATTTCTTAATCTTTGTGAAATATATGTTCAATGAGATTTTAAGTTGAGTG 2122
QY 83 TCTTAATGTAAGTACACGAGACTGATTCAGTCCGTGTACAATGCAACCAACCGCA 142
DB 2121 TCTTAATGTAAGTACACGAGACTGATTCAGTCCGTGTACAATGCAACCAACCGCA 2062
QY 143 ACTGTTACGAG-----AACAGAAATGCACTATATGATGATTTCAACCAATC 196
```

Db 2061 GTTCTGACAGAGGCAAGAGCGATGAGTAAGTAGGGTAGCAAGCCCAAGTC 2002
 Qy 197 TGTCAATATAGAAATTTAAACCAACATCCATGTTCTTAATTAAGGGTAATGCTTGGG 256
 Db 2001 CGTGCCATACAAATTTGGCCATCAGGCGCAAGGAGGCTTAAGCTGGGAATGATGCTCGCG 1942
 Qy 257 TTGGGGTGGAAACAAGGCTCAACCCGACCGGGGTTATTTGCTAACAGAGGGCAT 316
 Db 1941 TTGGGGTGGCAACAAGGCTCACTCTCAGCGCTGGATCTCGCCCAACAAGAGATAT 1882
 Qy 317 TTGATGGGCTACAAAGGA-CAGATTTCAACAGCAATTAATTGGCTCCCTCAAG 375
 Db 1881 CTATGGGGAGCAACAAGGATGGCGTTGAGCAGCGCAATTAATTCCCTCACTCAAG 1822
 Qy 376 CCTGACATCGAGTTGGGCTCTTC-----AGGAGAGAAATCTATGCCCAT 427
 Db 1821 CCTGACATCGGCGCGGCGCTCTCTTCAAGGTTGAAGGAAGATGATGACTTTTAA 1762
 Qy 428 CAAGAGCGCTGCTTCAATGGTTAAACCTGACGACATTTGTTGGGGGATGGATATCAG 487
 Db 1761 GAAGCAATTTCAACCGATGTCACACCCCAAGCATCTGATGATGGCGGTTGGATATCTC 1702
 Qy 488 CAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGAGTTGATGATCGATTTGGAGA 547
 Db 1701 CAACATGAACCTGGCGGATGCCATGGATGCTGCCCGGCTGCTGATGATCTACAGA 1642
 Qy 548 GCAATGAGCGCTTCAATGGAATCCATGCTTCACTCCCGGAATCTATGACCCGATTT 607
 Db 1641 GCAACTTCTGCTCTCATGAGATATGTTACCAATGCTGATCTACACCCCACTT 1582
 Qy 608 CATGCTGCAACCAAGAGAGGCTGCAACAAGTCAATCAAGGGCAACAAGCAAGCA 667
 Db 1581 CATGCTGCTACCAAGAGAGGCTGCTGCAACAATCTCATCAAGGGTTCCAAAGAGAGCA 1522
 Qy 668 AGTTCAACAATCATCAAGACATCAAGCGCTTTAAGAGCCCAAGAGTGAACAAGT 727
 Db 1521 AATGAGCGGATATATCAGAGCATACAGGAATTCMAAGGCGCAACAAGTTGACAAGT 1462
 Qy 728 GGTGATCTGTGATCTGCCAACAAGAGATCAAGTAATTTGGTTGGGCTTATGA 787
 Db 1461 AGTGTACTGTGACCGCTAACACAAGCGCTTACAGCATGATGTTGGGCTGATGA 1402
 Qy 788 CA-CCATGAGAAATCTTTGGCGCTGCTGAGCAGAANTGAGTGAATTTCTCTTCA 846
 Db 1401 CACCAAGAGAAATCTTTGATTTCTGAGAGAGAGAAATTCGCGCTCA 1342
 Qy 847 CCTGTATGCAATGCTTGTGTTATGAAATGTTCTTTCATTAATGAAGCCCTCA 906
 Db 1341 CCTGTACGCTCTCGCTTGATCCAGAGAGGTTCCCTTCAACAGGAGTCCCA 1282
 Qy 907 ACACTTTTGTACAGAGGCTGATGATCTTGCCATGCGAGAGAACCTTTGATTTG 966
 Db 1281 ACACATTTCTCTCGGAGTAATGAAATGACGCTTTGAAAGAACTATTAATTTG 1222
 Qy 967 ATGACTTCAAGAGTGTGACAGCAAAATGAAATCTGTTGTTGATTTCTTGTGGGG 1026
 Db 1221 ACAGACTTCAAGAGCGGTCAACGAAATGAAATCGTTTCAAGGACCTTTGAGG 1162
 Qy 1027 CTGATATCAAGCAACATCTATAGTCAATTAACCATCTGGGAAACAATGATGATGA 1086
 Db 1161 CGGGAATTAAGCAAGTCAATTTGATGCTACACATTTGGGCAACAAGATGATGA 1102
 Qy 1087 ATCTTTGCGCTCAACAACCTTTCTGTTCAAGAGAAATCTTCAAGAGCAAGTTGATG 1146
 Db 1101 ACTTTGCGCGCTCAACAACCTTTCTGTTCAAGAGAAATCTTCAAGAGCAAGTTGATG 1042
 Qy 1147 ATATGCTCAACAGCAATGCAATCTCTATAGAGCTGTAACATCAAGCAAGTTGATG 1206
 Db 1041 ACATGAGGTGACCAATGCAATCTATAGAGCGCGCGGTGAACATCCGACCAAGTTATG 982
 Qy 1207 TTATTAATGATGCTTATGATGAGGAGACAGCAAGAGACCATGATGATCACTTCA 1266
 Db 981 TGATCAAGTAAGTCTTATGTTGAGTGAACAGCAAGGCGCAATGACGATCACTTCCG 922

Qy 1267 AGATATTCATGAGGTGAAAGACACCATTTGTTGACAAACAGACGAGATTTCCCTC 1326
 Db 921 AGATCTTCATGAGCGGAGCGTACACATGCTATGACAAATAGTCCGAAGATTTTAC 862
 Qy 1327 TAGCTGCTCTATATTAATCTTGACCTTGGTCTTCTGAGCTCAGACATGAAATGAGT 1386
 Db 861 TGCCGCGCGCTCATCTTTGGACTTGGTCTTCTGAGCTCAGCTCAGCAAGATCCAT 802
 Qy 1387 TTAAAGCTGAAATGAGAAATTCACACTATTCACCCAGTTGCTACATCTCAGCT 1446
 Db 801 TCAGAGAGACGAGACAGAGAGTTCATTTCTTCAACCGGTTGCGTACTCTAGCT 742
 Qy 1447 ACTTCAACAAAGCTCTCTGTTTCCACCGGATACACATGATGGAATGATTTGCAAGC 1506
 Db 741 ACCTTCAACAAAGCTCTCTTGTGTAACCCAGGTAACCCAGTGTGTAACGGTGGTCAAGC 682
 Qy 1507 AGGTCATATGCTGTAACCAATATATAGAGGCTTGTGTAATGAGCCCAAGAAATACA 1566
 Db 681 AGAGGCGATGCTAGAGATATCATGCGCGCTGATGGGTTGTCCCGACATATACA 622
 Qy 1567 TGATTTCTGAGTACAGT 1584
 Db 621 TGATGCTAGATATCAAGT 604
 RESULT 5
 CK277950 988 bp mRNA linear EST 03-AUG-2004
 EST724028 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POAE302 5' end, mRNA sequence.
 CK277950
 CK277950.1 GI:39834928
 EST.
 Solanum tuberosum (potato)
 Solanum tuberosum
 Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;
 asterids: lamiales: Solanales: Solanaceae: Solanum.
 1 (bases 1 to 988)
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
 Generation of ESTs from abiotic stressed potato tissue
 Unpublished (2003)
 Other_ESTs: EST724029
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from the University of Arizona Genomics
 Institute via <http://genome.arizona.edu/orders/>.
 Seq primer: ATG TAG GTG ACA CTA TAG.
 Location/Qualifiers
 1..988
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POAE302"
 /tissue_type="abiotic stress treated leaf and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="potato abiotic stress cDNA library"
 /note="Vector: pCMVSPORT1; Site 1: BclRI; Site 2: NotI;
 supplier: Solanum tuberosum var. Kennebec plants were
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation
 of the soil with 150 mM NaCl and tissues were harvested at
 following application of the salt stress (leaves: 2hr,
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
 Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering
 applications. Drought stressed plants were harvested after
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d

and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Query Match 40.2%; Score 706.8; DB 7; Length 988;
Best Local Similarity 82.9%; Pred. No. 9e-194;
Matches 807; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

```

OY 230 TCTTAATTAGGGGTAATGCTTGTGGGTTGGGGTGAACAACGGCTCAACCTTCAACCG 289
DB 15 TTCCAATTAGGGGTTATGCTTGTGGATGGGGTGAACAATGGTTCAAGCTTGAAGCTGG 74
OY 290 TGGTGTATGCTTAACCGAGAGGCAATTCATGAGGCTCAACAAAGATCAACAAGC 349
DB 75 AGGTGTATGCTTAACAGAGAGATTCATGAGGCTCAACAAAGATCAACAAGC 134
OY 350 CAATTACTTGGGCTCCCTCACCAGGCTCAGCTATCCGAGTGGGTCCTTCCAGGGAGA 409
DB 135 CAATTACTTGGGCTCCTTCAACAGGCTCCTATGAGTGGTGGTCTTCAATGAGAGA 194
OY 410 GGAATCTATGCCCCCATTCAGAGCCTGCTTCATGTTAACTTCAAGCAGCATTTGTT 469
DB 195 GGAGATCTATGCCCCCTTAAAGACATCTCCATGCTCAATCCAGATGATGATGTT 254
OY 470 TGGGGATGAGGATATCAGACATGAACCTGGCTGATGCTAGGCGCAGGCGAAAGTGT 529
DB 255 TGGAGATGAGGACATGACATGATTTGGAGATGCTAGGCGCAGGCGCTAAAGTTT 314
OY 530 TGACATGCTTTCAGAGAGCAGTGGGCTTACATGATGATGATGCTTCACTCCCGG 589
DB 315 TGACATGCTTTCAGAGAGCAGTGGGCTTACATGATGATGATGCTTCACTCCCGG 374
OY 590 AATCTATGACCGGATTTGATGCTGCTCAACAGAGGCTGCCAACAAGCTCATCA 649
DB 375 TATCTATGACCGGATTTGATGCTGCTCAACAGAGGCTGCCAACAAGCTCATTA 434
OY 650 GGGCAAGAGAGGAGCAAGTTCACAAATCATCAAGATCAAGGCGTTTAAAGAAC 709
DB 435 AGGAACCAAGAGAGCAAGTTCACAAATCATCAAGATCAAGGCGTTTAAAGAAC 494
OY 710 CACCAAGTGAACAAGTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 769
DB 495 CACCAAGTGAACAAGTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 554
OY 770 GGTGTGGGCTTATGACACATGAGAACTCTTGGCTGCTGTGAGACAGAAATGAGC 829
DB 555 GGCTGTGGCTTATGATGATGAGAAACCTCTTGGCTGCTGTGAGATGAGAAATGAGC 614
OY 830 TGAAGATTTCTTCACCTGATGATGATGATGATGATGATGATGATGATGATGAT 889
DB 615 TGAAGATTTCTTCACCTGATGATGATGATGATGATGATGATGATGATGATGAT 674
OY 890 TATGAGACCCCTCAGAACATTTGTACAGAGGCTGATGATCTTGCATCGCAGAGAA 949
DB 675 CATGGAACCCCAACAACAATTTGTCCAGGCTCATGATTTGGCATCAAGAGAGAA 734
OY 950 CACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1009
DB 735 CACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 794
OY 1010 TATTTTCTTGTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1069
DB 795 TATTTTCTTGTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 854

```

```

OY 1070 AACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1129
DB 855 TAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 914
OY 1130 GAGCACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1189
DB 915 AAGCAATTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 974
OY 1190 TCCAGACCATGTTG 1203
DB 975 CCTGACCATGTTG 988

```

RESULT 6

AW348136/c 782 bp mRNA linear EST 04-OCT-2000
LOCUS 1001001A21B10 Gm-r1021 Glycine max cDNA clone Gm-r1021-68 3', mRNA
DEFINITION

ACCESSION

AW348136
AW348136.1 GI:6845846
EST.

KEYWORDS

Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE

1 (bases 1 to 782)
Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,
Erpelting, J., Ragh, C., Shoop, B., Pardinas, J., Liu, L. and Lewin, H.,
A Functional Genomics Program for Soybean (NSF 9872565)
Unpublished (1999)
Other ESTs: A1442485

AUTHORS

Contact: Vodkin, L.O., PI, A Functional Genomics Program for
Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional
Genomics

JOURNAL

University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582

COMMENT

Email: l-vodkin@uiuc.edu
This clone is available through: Genome Systems, Inc. 4613 World
Parkway Circle St. Louis, Missouri 63134. For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or info@genome
systems.com web site: www.genomesystems.com

FEATURES

Location/Qualifiers
1..782
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="Gm-r1021-68"
/tissue_type="root"
/lab_host="X110-Gold"
/clone_lib="Gm-r1021"
/note="Vector: pBluescript II Xr; Site 1: EcoRI; Site 2:
XhoI; Library Gm-r1021 is a sequence-driven, retracted set
of the original library Gm-c1004 which was prepared from
root cDNA. The mRNA was isolated from entire roots of 8
day old 'Williams' seedlings which were propagated on
paper towels with distilled water. Stratagene's cDNA
Synthesis Kit (catalog #200401) was used to synthesize
the cDNA. The Gm-c1004 library was constructed by Dr.
Paul Keim & Virginia H. Coryell, Department of Biology,
Box5640, Northern Arizona University, Flagstaff, AZ
86011, email: paul.keim@uau.edu, virginia.coryell@uau.edu.
The cDNA analysis to select unique genes was performed
by the laboratory of Ernest Retzel, Computational Biology
Centers, University of Minnesota,
http://www.cbc.umn.edu/researchprojects/soybean/index.html

Db	754	AAATGAATCTGTGTGGTTGANNNNNTTGTGGGGCTGGGATCAAGCCNNNATCTATAG	695
QY	1051	TCAGTTACACCACTCTGGGAAACAATGATGTATGATCTTTGGCTCCACAACTTTC	1110
Db	694	TTAGTTACAAACCATCTGGGAAACAATGTGTATGAAATCTCGGCTCCAAACCTTCC	635
QY	1111	GTTCCAAGGAATCTCCCAAGGCAACGTTGTATGATATATGTCAACGCAATGCATCC	1170
Db	634	GCTCCAAAGAAATCTCCAAAGCAACGTTGTAGATATGTCAACGCAATGCATCC	575
QY	1171	TCTATGAGCCCTGTGTGACATCCAGACCATGTTGTATTATTAAGTATGTGCTTAAGT	1230
Db	574	TCTATGAGCCCTGTGTGAAATCCCGACCATGTTGTATTATTAAGTATGTGCTTAAGT	515
QY	1231	GGGACAGCAAGAGCCATGTGATGATGACACTTCAGAGATATTCATGGGTGGAAAGACA	1290
Db	514	GGGATAGCAAGAGCCATGTGATGATGACACTTCAGAGATATTCATGGGTGGAAAGACA	455
QY	1291	CCATTGTTTGCACAAACATGCGAGATCCCTTGTGCTGCTCTATTATCTTGGACT	1350
Db	454	CCATTGTTTGCACAAACATGTGAGATTCCTTTTATGCTGCTCTATTATCTTGGACT	395
QY	1351	TGTCCTCTTCTGTGAGCTCAGCACTAGAAATCGAGTTTAAAGCTGAAATGAGGAAAT	1410
Db	394	TGTCCTCTTCTGTGAGCTGAGCACTAGAAATCCAGTTTAAAGCTGAAATGAGGAAAT	335
QY	1411	TTCACCTCATCCACCAGTGTGACATCCCTCAGTACTCTACCAAGGCTCCTCGGTC	1470
Db	334	TTCACCTCATCCACCAGTGTGACATCTCAGTACTCTACCAAGGCTCCTCGGTC	275
QY	1471	CACGGGATACACAGTGTGATGATGTCAAAAGCAGCGTGCATCGCTGGAAAACTAA	1530
Db	274	CACGGGATACACAGTGTGATGATGTCAAAAGCAGCGTGCATCGCTGGAAAACTAA	215
QY	1531	TGAGGCTTGTGTGAGTTGGCCCCAGAAATACATGATTTCTCGAGTACAAAGTGAGCA	1590
Db	214	TGAGGCTTGTGTGAGTTGGCCCCAGAAATACATGATTTCTCGAGTACAAAGTGAGCA	155
QY	1591	TGGGACCAAGAAATATATAGTGGGGATGAGCTGATGATGTTTATGTTAAATATATG	1650
Db	154	TGGGATCAAAATATATATTTGTGGGGATGAGCTGATGATGTTTATGTTAAATATATG	95
QY	1651	TTTGTCTTAAATTTTGCAGGTGTAAATGATGATCACTTCATTTATGCTTTAGACGG	1710
Db	94	TTTGTCTTAAATTTTGCAGGTGTAAATGATGATCACTTC-TCATTAATGCTTTAGACAG	36
QY	1711	GGCATATTCGTTTACTAGGAACATGATGAT 1743	
Db	35	GACATATTCGTTTCTAGGACATGAATGAT 3	

FEATURES

Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics
University of Illinois
Ewain R. Malignan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vock@uiuc.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134. For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contract: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
Seq Primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.

Source

```

/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/cclone="Gm-r1021-3671"
/rissue_type="root"
/lab_host="XLI0-Go1d"
/cclone_lib="Gm-r1021"
/notes="Vector: pBluescript II XR, Site 1: EcoRI; Site 2:
XhoI; Library Gm-r1021 is a sequence-driven, reracked set
of the original library Gm-cl004 which was prepared from
root cDNA. The mRNA was isolated from entire roots of 8
day old 'Williams' seedlings which were propagated on
paper towels with distilled water. Stragane's cDNA
Synthesis Kit (catelogs #200401) was used to synthesize
the cDNA. The Gm-cl004 library was constructed by Dr.
Paul Keim & Virginia H. Coryell, Department of Biology,
Box56640, Northern Arizona University, Flagstaff, AZ
86011, email: paul.keim@nau.edu, virginia.coryell@nau.edu.
The contig analysis to select unique genes was performed
by the laboratory of Ernest Retzel, Computational Biology
Centers, University of Minnesota,
http://www.abc.umn.edu/Research/Projects/Soybean/index.html
. Reracking was performed by Genome Systems, St. Louis,
http://www.genomesystems.com, and sequencing by the Keck
Center for Comparative and Functional Genomics,
University of Illinois,
http://www.illc.uiuc.edu/biotech/keck.html."

```

ORIGIN

Query Match	37.4%;	Score 658.4;	DB 2;	Length 686;
Best Local Similarity	96.1%;	Pred. No. 9.7e-180;		
Matches 659;	Conservative 0;	Mismatches 27;	Indels 0;	Gaps 0;

[illegible]

Db	686	TAGTCAGTTNNA	CCCA	NNNNNNNNNN	NTA	TGANGATATNN	TCCTTCGGCTCC	CAAACTT	627
Qy	1108	TCGCGTCCAA	GGAAT	CTCC	CAAGCAAC	GTGTGATGATAT	GTGCAAC	GCAATGCA	1167
Db	626	TCGCGTCCNN	GNNA	TCTCC	CAAGCAAC	GGNNGTTGA	TGATATGTCA	CAACGAATGCA	567
Qy	1168	TCCTCTATGA	CCCTG	TGAACAT	CCAGACCA	TGTGTGTTATTA	TAA	GATATGCTTACG	1227
Db	566	TNNCTATGAC	CCGTGTGA	CATCC	CAACCAT	GTGTGTTATTA	TAA	GATATGCTTACG	507
Qy	1228	TAGGGGACAG	CAAGAGCC	ATG	TGATGAT	CACTTCAGAGAT	ATTCAT	TGGGTGAAAGA	1287
Db	506	TAGGGGACAG	CAAGAGCC	ATG	TGATGAT	CACTTCAGAGAT	ATTCAT	TGGGTGAAAGA	447
Qy	1288	GCACCATTTG	TTCACAC	ACATG	CGAGATTC	CCCTCTTAG	CGTCTCTTA	ATCTTGG	1347
Db	446	GCACCATTTG	TTCACAC	ACATG	CGAGATTC	CCCTCTTAG	CGTCTCTTA	ATCTTGG	387
Qy	1348	ACTTGGTCTT	CTTCTG	TAGCTC	AGCATAG	AATGAGTTAA	AACTG	GAATAAGGGAA	1407
Db	386	ACTTGGTCTT	CTTCTG	TAGCTC	AGCATAG	AATGAGTTAA	AACTG	GAATAAGGGAA	327

QY	1378	GAATCGAGTTTAAAGCTGAAAATG--AGGGAAAATTCATCATTCACCCAGTTGGTAC	1435
Db	841	GCATTACAGCTAAAGAGCTGAAAGGGAGGGGAAAGTTCCATCTTCACCTGTGTCTAC	900
QY	1436	CATCCTCAGCTACCTGCACCAAGGCTCCTCTGTGTTCC	1471
Db	901	TATTTCTAGCTACCTGCACCAAGGCTCCCTGTGTACC	936
RESULT 10			
LOCUS	CK275352		
DEFINITION	CK275352	914 bp	linear EST 03-AUG-2004
	EST721430 potato abiotic stress cDNA library		Solanum tuberosum cDNA
ACCESSION	CK275352		
VERSION	CK275352		
KEYWORDS	CK275352.1	GI:39832330	
SOURCE	EST.		
ORGANISM	Solanum tuberosum		
	Solanum tuberosum (potato)		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;		
	asterids; lamids; Solanales; Solanales; Solanales; Solanum.		
REFERENCE	1 (bases 1 to 914)		
AUTHORS	Buell, C.R., Hart, A., Zismann, V., Karamcheva, S.A. and Baker, B.		
TITLE	Generation of ESTs from abiotic stressed potato tissue		
JOURNAL	Unpublished (2003)		
COMMENT	Contact: Robin Buell		

Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES	Location/Qualifiers
source	1. .914

`/organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:1113"
 /clone="POADM85"
 /tissue_type="abiotic stress treated leaf and root tissue"
 /lab_host="DH108-TonA"
 /clone_lib="potato abiotic stress cDNA library"
 /note="vector: pCMVSPORT6.1, Site 1: EcoRI, Site 2: NotI;
 supplier: Solanum tuberosum var. Kennebec plants were
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation of
 the soil with 150 mM NaCl and tissues were harvested at
 following application of the salt stress (leaves: 2hr,
 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d) .
 Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering
 applications. Drought stressed plants were harvested after
 cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
 and 5d) . Set 3 were grown under the standard conditions
 and then were cold stressed by placement at 4 C. Cold
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d. Set 4 were grown under the standard conditions and
 then were heat stressed by placement at 35 C. Heat
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d and 4d and heat-stressed roots were harvested at 6 hr,
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and
 equal RNA from each tissue and stress was pooled to
 construct the cDNA library. RNA sample."`

ORIGIN

Query Match	37.1%;	Score 653.4;	DB 7;	Length 914;
Best Local Similarity	83.5%;	Pred. No. 3.1e-178;		
Matches 741;	Conservative 0;	Mismatches 146;	Indels 0;	Gaps 0;

702 AAGGAGCCACCAAGTGGACAAGTGGTTGTACTGTGGACTGCCAACACAGAGAGGTAC 761

Db	1	AAGGAGAAACAGCAAAAGTGGACAAAGTGCTGGTGTATTGTGACTGCCAACACTGAAAGATAC	60
OY	762	AGTAATTTGGTTGTGGGCTTAATGACCACTGGAGAAATCTTTGGCTGCTGTGGACAGA	821
Db	61	AGCAGTGTGGCTGTGGCTTAATGAATACATGAGAAAACCTTTGGCTTGTGTGAATAGA	120
OY	822	AATAGGCTGAGATTTCTCCCTCCACCTGTGATGCCATGCTGTGTGTTATGGAATAATGTT	881
Db	121	AATAGGCGGAATATCTCTTCCACCTGTATGCTATTGCTGTATTTCTTGAAAAATGTG	180
OY	882	CCTTTCATTAATGAGAGCCCTCAGAACACTTTTGTATACAGGGCCTAATGATCTTGCCATC	941
Db	181	CCCTTATTCATGAGAACGCCAACAACACTTTGTGCCAGGGCCTCATTTGATTTGGCCATC	240
OY	942	GCGAGAAACACTTTGATTTGGTGGAGATGACTTCAAGAGTGTGAGACCAAAATGAATCT	1001
Db	241	AAGAGGAACACTTTGATTTGGTGGAGATGACTTAAAGTGTGTCAACCAAGATATAGTCA	300
OY	1002	GTTGTTGGTGAATTCCTTTGGGGGGCTGTATCAAGGCAACATATATAGTCAGTTCAAC	1061
Db	301	GTGCTGGTGAATTTCTTTGTTGGAGCTGTATTAAGCCACATATATTTGAGCTTCAAC	360
OY	1062	CATCTGGGAAACAAATGATGATGATATCTTTGGCTCCACAACTTTCCGTTCCAGAGAA	1121
Db	361	CATTTGGGTAAACATGATGATGATGATTAATCTGTGGCCCTCAGACTTTCCGTTCAAAGAG	420
OY	1122	ATCTTCAAGACAAAGTTGTGATGATATGTGTCAACAGCAATGCCATCTCTATAGAGCT	1181
Db	421	ATCTGAAAGCAATGTTGTGATGATGATGATGTTCTGTATGTCATCTTTATGAGGCT	480
OY	1182	GCTAACAATCCAGACCAATGTTGTGTTATTAAGTATGTCCTTAAGAGGGGACAGCAG	1241
Db	481	GGAAGACCCCTGACCAATGTTGTGTTATTAAGTATGATACCAATATGTGGAGATAGCAG	540
OY	1242	AGAGCCATGATGATGATCACTTCAAGATATATTCAGGGGTGGAAAAAGACCACTTGTGTTG	1301
Db	541	AGGCGATGATGATGATGATCACTCAAGATTTTCAATGGGTGGAAAGAACCACTTGTATG	600
OY	1302	CACAACAATGCGAGGATTCCTTCTTATGCTGCTCTATATCTTGAGCTTGTGCTTCTT	1361
Db	601	CACAATACCTGTGAGGATTCCTTTTGGCTGCTCAATATATCTTGGATTGTGCTCTCTTG	660
OY	1362	GCTAAGCTCAGACTAAGATGATGATTTAAAGCTBAAAAATGAGGGAAAAATCCACTCAATC	1421
Db	661	GCTAATCTCAGACCCCGATTCACCTCAAGCTBAAGAGAGGGTAAATGTTCACTCTCTTC	720
OY	1422	CACCAAGTTGTACATCTCAGCTACCTCAACAAGCTCCTCTGTGTTCCACCGGGTACA	1481
Db	721	CATCCCGTGGGAGACATCTCTCAGCTATCTTAACAAGGCTCCTCTGTATCCACAGATACA	780
OY	1482	CCAGTGTGAATGATGTCTCAAGACGCTGTCAAATGCTGAAAAATATATAGGGCTTGT	1541
Db	781	CCAGTGTGAATGCTCTTCAAAACAGAGGCAATGCTTGAGAAATATATGAGGGCTGTGT	840
OY	1542	GTTGATTTGGCCCAAGAAATTAATGATTTCTCAGATACAAGTAAAG	1588
Db	841	GTTGACTGTGGACACAGAAACAAATGATATTTGAAAAATACATATAAG	887

RESULT 11			
LOCUS	672 bp	mRNA	linear
DEFINITION	EST 03-DEC-2001		
AM508269	AM508269	g:1h10.y1 Gm-ri030 Glycine max cDNA clone	GENOME SYSTEMS CLONE ID: 1
		ri030-2492 5' similar to SM:INOL_C1RPA P42802	
		MYO-INOSITOL-1-PHOSPHATE SYNTHASE ; mRNA sequence.	

ACCESSION AW508269

KEYWORDS	EST.
SOURCE	glycine max (soybean)
ORGANISM	glycine max

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

CDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGTGTATCAAGCAGAGTGGCAGCGG-3' and 5'-ATTCTAGAGCCGAGCGGCGCAGATG-dT(30)-NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Query Match 36.8%; Score 647.4; DB 6; Length 867;
Best Local Similarity 84.2%; Pred. No. 1.7e-176;
Matches 729; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```

QY 108 GAGATTCACTCCGTGTACACTACGAAACCCGAACTTTGTTCAGAGAACGAGATGGC 167
DB 1 GAGATTCACTCCGTGTACACTACGAAACCCGAACTTTGTTCAGAGAACGAGATGGC 60
QY 168 ACCTATCAGTGGATGTGTCAACCCAAATGTCTCAATTCGATTTTAAACCAATCCAT 227
DB 61 ACCTATCAGTGGATGTGTCAACCCAAATGTCTCAATTCGATTTTAAACCAATCCAT 120
QY 228 GTTCTTAATTTAGGGGTATGCTTGTGGGTGGGTAACAAGCGCTCAACCTCACC 287
DB 121 GTCCCATGCTAGGGGTATGCTTGTGGGTGGGTAACAAGCGCTCATCTCTCACT 180
QY 288 GGTGTGTATTGCTTACCGAGAGGGCATTTATGCGCTTCAAGAGACAGATTCAACA 347
DB 181 GGTGTGTATTGCTTACCGAGAGGGCATTTATGCGCTTCAAGAGACAGATTCAACA 240
QY 348 GCATTTACTTTGGGCTCCCTCAACCAAGCTCAGCATTCGAGTTGGGCTTTCCAGGGA 407
DB 241 GCATTTACTTTGGGCTCCCTCAACCAAGCTCAGCATTCGAGTTGGGCTTTCCAGGGA 300
QY 408 GAGAAATCTATGCCCATTTCAAGAGCTCTTCCAAATGGTTTACCTTGACGACATTTGT 467
DB 301 GAGAGATTTAAGCTCCATTTCAAGAGCTCTCTCCATGGTAAACCAATGACATTTGT 360
QY 468 TTTGGGGGATGGGATATCGACAACATGAACCTGGCTGATGCCATGGCCAGGCAAAAGTG 527
DB 361 TTTGGGGGATGGGATATCGACAACATGAACCTGGCTGATGCCATGGCCAGGCTTAAAGTTC 420
QY 528 TTTGATCATGATTTTGAGAGAGATTTGAGGCTTACATGGAATCCATGCTTCACTCCC 587
DB 421 CTGGACATGATCTGAGAGAGATTTGAGGCTTACATGGAATCCATGCTTCACTCCC 480
QY 588 GGAATCTATGACCCGATTTTCAATGCTGCCCAACCAAGAGAGCTGGCCAAACGTCATC 647
DB 481 GGAATCTATGACCCGATTTTCAATGCTGCCCAACCAAGAGAGCTGGCCAAACGTCATC 540
QY 648 AAGGCAACAAGCAAGAGCAAGTTCACAATCATCAAGACATCAAGCGCTTTAAGGA 707
DB 541 AAGGCAACAAGCAAGAGCAAGTTCCAGCATTTAAAGATATTAGGAGTTTAAAGAG 600
QY 708 GGCACCAAGTGGAGCAAGGTGTGTGATCTGTGACCTGCCAACAAGAGATCAAGTAT 767
DB 601 GGCACCAAGTGGAGTGGTGTGTGATCTGTGACCTGCCAACAAGAGATCAAGTAT 660
QY 768 TTGTTTGGGGCTTATGACACATGAGAACTCTTGGCTGCTGTGAGCAAGAAATGAG 827
DB 661 TTGTTTGGGGCTTATGACACATGAGAAATCTTGGCTGCTGTGAGCAAGAAATGAG 720
QY 828 GCTGAGATTTCTCTTCCACCTTTGATGCAATGCTTGTGTTAGGAAATGTTCTTTTC 887
DB 721 CCTGAGATTTCTCTTCCACCTTTGATGCTTGTGTTAGGAAATGTTCTTTTC 780
QY 888 ATTATGGAAGCCCTCAGAACCTTTTGAACAAGGCTGATGATCTTGGCATTCGCGAG 947
DB 781 ATTCAATGAGAGCCCAAGAACCTTTTGTTCAGAGGCTGATGATTTGCTATTAAAGAG 840
QY 948 AACACTTTGATGTTGAGATGACTT 973
DB 841 AACACTTTGATGTTGAGATGACTT 866

```

RESULT 13
CO085839 865 bp mRNA linear EST 16-JUN-2004
LOCUS GR_Ea03B19.f GR_Ea Gossypium raimondii cDNA clone GR_Ea03B19 5',
DEFINITION mRNA sequence.

ACCESSION CO085839
VERSION CO085839.1 GI:48776473
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustoids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE 1 (bases 1 to 865)
Kim,H., Yu,Y., Kudrna,D., Hatfield,D., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 03 row: E column: 19.

FEATURES
source
1..865
location/Qualifiers
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea03B19"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_lib="GR_Ea"
/note="vector: PCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 36.5%; Score 642.6; DB 7; Length 865;
Best Local Similarity 83.9%; Pred. No. 4.2e-175;
Matches 726; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

```

QY 479 GATATCGCAACATGAACCTGGCTGATGCAATGGCCAGGGCAAAAGTGTGACATCGA 538
DB 1 GATATTTAGTACATGAACCTGAGTCTATGGCTTAAAGGCTTTTCGACATGGA 60
QY 539 TTTGCAAGAGAGTTGAGGCTTACATGGAATCCATGCTTCCACTCCCGGAATCTATGA 598
DB 61 TCTGCAAAAGCACTGAAACCTTACATGGAATTCATGATCCCATCTCCGTAATTTAGA 120
QY 599 CCCGATTTTCAATGCTGCCCAACAAGAGAGCTGCCCAACAGTCATCAAGGCAAAA 658
DB 121 TCTGATTTCAATGCTGCTTAAACAAGGTGAACGTGCCAATATGTCATCAAGGGAACAA 180
QY 659 GCAAGAGCAAGTTCAACAATCATCAAGAATCAAGGCGTTTAAAGGAAGCAACCAAGT 718
DB 181 GAAAGAACAAAGTTTCAGCGATCATCAAGACATCAAGAGATTCAAGGAAAGAAACCAAGT 240
QY 719 GGACAAAGTGTGTGATCTGTGACCTGCCAACAAGAGATCAAGTAAATTTGTTGTTGGG 778
DB 241 GGACAAAGTGTGTGATCTGTGACCTGCCAACAAGAGATCAAGTAAATTTGTTGTTGGG 300
QY 779 CTTTAAATGACACATGAGAAATCTTTGGCTGCTGTGTGACGAAGAAATGAGGCTGAGATTTC 838
DB 301 GCTAAATGACACCGTGGAAGCCTTATGCTTCTTTGGAAGAAATGAAATCAGAGATTTC 360
QY 839 TCTTCCACCTTGTATGCAATGCTTGTGTTATGGAAGAAATGTTCTTCAATTAATGAGAG 898

```

Db 361 TCCTTCACCTTGTATGATGCTATGCTGTGTCTTGAAGAAAGTTCCTTCATCAATGAGC 420
QY 899 CCTTCAGAACCTTTTGTACACAGGCTGATTTGATCTTCCATGCGAGAAACCTTTGAT 958
Db 421 CCCACAAAACCTTTGTTCACAGGCTGATTTGATTTGGCTATTCAGAGAACTGTCTGAT 480
QY 959 TGTGAGATGATCTCAAGAGTGTGAGACCAAAATGAATCTGTGTTGATTTTCT 1018
Db 481 TGAAGAGATGATCTCAAGAGTGTGAGACCAAAATGAATCTGTGTTGATTTTCT 540
QY 1019 TGTGAGGAGCTGTATGATCAAGCAATCTATGATCAAGCAATCTGTGTTGATTTTCT 1078
Db 541 TGTGAGGAGCTGTATGATCAAGCAATCTATGATCAAGCAATCTGTGTTGATTTTCT 600
QY 1079 TGTGATGAATCTTTGCGCTTCACAACTTTCCGTTCCAAAGAAATCTCCAAAGCAAGT 1138
Db 601 TGTGATGAATCTTTGCGCTTCACAACTTTCCGTTCCAAAGAAATCTCCAAAGCAAGT 660
QY 1139 TGTGATGAATCTTTGCGCTTCACAACTTTCCGTTCCAAAGAAATCTCCAAAGCAAGT 1198
Db 661 TGTGATGAATCTTTGCGCTTCACAACTTTCCGTTCCAAAGAAATCTCCAAAGCAAGT 720
QY 1199 TGTGATGAATCTTTGCGCTTCACAACTTTCCGTTCCAAAGAAATCTCCAAAGCAAGT 1258
Db 721 TGTGATGAATCTTTGCGCTTCACAACTTTCCGTTCCAAAGAAATCTCCAAAGCAAGT 780
QY 1259 CACTTCAGATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1318
Db 781 CACTTCAGATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 1319 TTCCCTTAGCTGCTCTATTTATC 1343
Db 841 TTCCCTTAGCTGCTCTATTTATC 865

RESULT 14
LOCUS COL13870 898 bp mRNA linear EST 16-JUN-2004
DEFINITION GR_Eb014121.r GR_Eb Gossypium raimondii cDNA clone GR_Eb014121
3', mRNA sequence.
ACCESSION COL13870.1 GI:48812557
VERSION COL13870.1
KEYWORDS EST.
SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 898)
Kim H., Yu Y., Kudrna D., Hatfield J., Stum D., Mueller C.,
Wing R.A., Rapp R.A., Wendel J.F., Rao K., Soderlund C. and
Wing R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 014 row: I column: 21.
Location/Qualifiers
1..898
FEATURES
Source
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb014121"
/issue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_1b="GR_Eb"
/note="Vector: pCMV.SORT-6.1; Site_1: NotI; Site_2:

ECORY: Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

Query Match 36.3%; Score 639.4; DB 7; Length 898;
Best Local Similarity 82.0%; Pred. No. 3.6e-174;
Matches 736; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

QY 198 GTCAATATGAAATTTAAACCAATCCATGTTCTTAATTAAGGGTAATGCTTGAGT 257
Db 1 GTCAATATGAAATTTAAACCAATCCATGTTCTTAATTAAGGGTAATGCTTGAGT 60
QY 258 TGGGTGGAACCAAGGCTCAACCTTCACCGGTGTGTATGCTTACCAAGGAGGAT 317
Db 61 TGGGAGGAAACAAATGTTCAACCTTCACCGGTGTGTATGCTTACCAAGGAGGAT 120
QY 318 TCATGGGCTACAAAGGACCAAGATTCACCAAGGCAATTAATGCTTCCCTCAAGCC 377
Db 121 TCTTGGGCTACAAAGGACCAAGATTCACCAAGGCAATTAATGCTTCCCTCAAGCC 180
QY 378 TCAGCTATCCAGTGTGGTCTTCAGAGGAGAAATCTATGCCCCATTCAGAGCTG 437
Db 181 TCAGCTATCCAGTGTGGTCTTCAGAGGAGAAATCTATGCCCCATTCAGAGCTG 240
QY 438 CTTCCAAATGTTAACTTCAGAGCAATGTTTGGGGATGGGATATACCAATGAC 497
Db 241 CTTCCAAATGTTAACTTCAGAGCAATGTTTGGGGATGGGATATACCAATGAC 300
QY 498 CTGGCTGATGCCATGGGCAAGGAGGTTTGAATCATGATTTGAGAAAGAGTTGAG 557
Db 301 CTAGCTGATGCCATGGGCAAGGAGGTTTGAATCATGATTTGAGAAAGAGTTGAG 360
QY 558 CCTTACATGAAATTCATGCTTCCACTCCCGGAATCTATGACCCGATTTCAATGCTG 617
Db 361 CCTTACATGAAATTCATGCTTCCACTCCCGGAATCTATGACCCGATTTCAATGCTG 420
QY 618 AACCAAGAGAGCGTCCCAACAGCTCATCAAGGGCAACCAAGAGAGCAATTCACAA 677
Db 421 AACCAAGAGAGCGTCCCAACAGCTCATCAAGGGCAACCAAGAGAGCAATTCACAA 480
QY 678 ATCATCAAAAGACATCAAGGCGTTTAAGAGCCACCAAGTGAACAGGTGTTGACTG 737
Db 481 ATCATCAAAAGACATCAAGGCGTTTAAGAGCCACCAAGTGAACAGGTGTTGACTG 540
QY 738 TGAATGCAACACAGAGAGTACAGTATTTGTTGGGCTTAATGACCAATGAG 797
Db 541 TGAATGCAACACAGAGAGTACAGTATTTGTTGGGCTTAATGACCAATGAG 600
QY 798 AATCTCTGCTGCTGTGAGCAAAATGAGCTGATTTCTCTTCCACTTGTATGCC 857
Db 601 AATCTCTGCTGCTGTGAGCAAAATGAGCTGATTTCTCTTCCACTTGTATGCC 660
QY 858 ATGCTGTGTTATGAAATGTTCTCTTCAATTAATGAGAGCCCTCAGAACCTTTGTA 917
Db 661 ATGCTGTGTTATGAAATGTTCTCTTCAATTAATGAGAGCCCTCAGAACCTTTGTA 720
QY 918 CCAAGGCTGATTTGATCTTCCATTCGAGAGAACCTTTGATTTGATGAGATCACTTCAAG 977
Db 721 CCAAGGCTGATTTGATCTTCCATTCGAGAGAACCTTTGATTTGATGAGATCACTTCAAG 780
QY 978 AGTGTGACACCAAAATGAATCTGTGTTGATTTCTTGTGAGGAGCTGTATCAAG 1037
Db 781 AGTGTGACACCAAAATGAATCTGTGTTGATTTCTTGTGAGGAGCTGTATCAAG 840
QY 1038 CCAACATCTATGATGATTAACACCATCTGGGAAACATGATGATGATCTTCCG 1095
Db 841 CCAACATCTATGATGATTAACACCATCTGGGAAATGATGATGATCTTCCG 898

RESULT 15
LOCUS CK320178 906 bp mRNA linear EST 11-MAY-2004

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 16:00:16 ; Search time 968.454 Seconds
(without alignments)
10758.117 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760

Sequence: 1 cctctcttatcctcttgcgt.....aatgtagtataatttgcgt 1760

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2002s:*
6: geneseqn2003s:*
7: geneseqn2004s:*
8: geneseqn2005s:*
9: geneseqn2006s:*
10: geneseqn2007s:*
11: geneseqn2008s:*
12: geneseqn2009s:*
13: geneseqn2010s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1760	100.0	1760	ADQ14490	Adq14490 Wild type
2	1760	100.0	1782	AAV62440	AAV62440 Soybean w
3	1533	87.1	1533	ADQ14498	Adq14498 Mutant m
4	1533	87.1	1533	ADQ14499	Adq14499 Soybean m
5	1533	87.1	1533	ADQ14500	Adq14500 Soybean m
6	1531.4	87.0	1533	AAV62443	AAV62443 Soybean m
7	1531.4	87.0	1533	ADQ14494	Adq14494 Mutant m
8	1531.4	87.0	1533	ADQ14495	Adq14495 Mutant m
9	1465.8	83.3	1533	ADQ14504	Adq14504 Wild type
10	1465.8	83.3	1533	ADQ14505	Adq14505 Mutant m
11	1465.8	83.3	1533	ADQ14506	Adq14506 Mutant m
12	1465.8	83.3	1533	ADQ14507	Adq14507 Mutant m
13	1464.2	83.2	1533	ADQ14508	Adq14508 Mutant m
14	1464.2	83.2	1533	ADQ14509	Adq14509 Mutant m
15	1120.4	63.7	1950	AAV90402	AAV90402 Nicotiana
16	1001.4	56.9	1781	AAV87643	AAV87643 Brassica
17	997.6	56.7	1938	AAV50242	AAV50242 Arabidops
18	995.4	56.6	1533	AB213633	Ab213633 Arabidops
19	995.4	56.6	1533	ADN73524	Adn73524 Thale cre
20	994.6	56.5	1837	AAV34806	AAV34806 Arabidops

21	994.2	56.5	1959	4	AAV85922	AAV85922 MIP synth
22	990.8	56.3	1665	2	AAV09006	AAV09006 Inducible
23	990.4	56.3	1536	3	AAV34328	AAV34328 Arabidops
24	984	55.9	1931	2	AAV24407	AAV24407 Maize myo
25	956.4	54.3	1759	6	ABQ72653	ABQ72653 Human MDD
26	835.8	47.5	1536	12	ADP43918	ADP43918 P. coarct
27	789.4	44.9	1719	3	AAV1095	AAV1095 Arabidops
28	552	31.4	1605	8	ABT19345	ABT19345 Aspergill
29	552	31.4	1605	8	ABT21165	ABT21165 Aspergill
30	514.4	29.2	2280	4	ABU17361	ABU17361 Drosophill
31	502.2	28.5	1772	4	AAH14171	AAH14171 Human CDN
32	502.2	28.5	1818	6	AAH15251	AAH15251 Human CDN
33	502.2	28.5	1825	6	ABT07176	ABT07176 Human ova
34	502.2	28.5	1852	4	AAH47741	AAH47741 Human hmt
35	502.2	28.5	1862	4	AAV59232	AAV59232 Human pol
36	502.2	28.5	2380	13	ADQ86789	ADQ86789 Human tum
37	502.2	28.5	2380	13	ACN085720	ACN085720 Human tum
38	502.2	28.5	2380	13	ACN40520	ACN40520 Human tum
39	487.2	27.7	1833	4	AAH47740	AAH47740 Murine m
40	485	27.6	1991	12	ADG73735	ADG73735 Aspergill
41	472.4	26.8	618	13	ACN62890	ACN62890 Cotton ca
42	456.8	26.0	1835	12	ADQ25429	ADQ25429 Human sof
43	454.4	25.8	1702	10	ADB62765	ADB62765 Human sof
44	451.2	25.6	2338	4	AAV10108	AAV10108 Human pol
45	441	25.1	1704	8	ABT20567	ABT20567 Aspergill

ALIGNMENTS

RESULT 1	ADQ14490	ADQ14490 standard; cDNA; 1760 BP.
XX	ADQ14490;	
XX	23-SEP-2004 (first entry)	
DT	XX	Wild type soybean myo-inositol 1-phosphate synthase cDNA #1.
DE	XX	Soybean, myo-inositol 1-phosphate synthase; gene; ss;
KW	XX	myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KM	XX	inosinic phosphate.
XX	XX	Glycine max.
OS	XX	
FT	FT	Key Location/Qualifiers
FT	FT	CDS 54..1586
FT	FT	/*tag= a
FT	FT	/product= "wild type soybean myo-inositol 1-phosphate
FT	FT	synthase #1"
PN	XX	US2004128713-A1.
PD	XX	01-JUL-2004.
XX	XX	
XX	XX	21-NOV-2003; 2003US-00718952.
XX	XX	
XX	XX	08-APR-1997; 97US-00835751.
XX	XX	07-APR-1998; 98WO-US0006822.
XX	XX	26-APR-1999; 99US-00289315.
XX	XX	11-MAR-2002; 2002US-00025003.
XX	XX	
XX	XX	(HITZ/) HITZ W D.
XX	XX	(SEBA/) SEBASTIAN S A.
XX	XX	(GRAC/) GRACE D J.
XX	XX	(STRE/) STREIT L G.
XX	XX	
XX	XX	Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX	XX	WPI; 2004-533135/51.
XX	XX	P-PSDB; ADQ14491.

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.

XX Claim 4; SEQ ID NO 1; 48pp; English.

XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.

XX Sequence 1760 BP; 494 A; 371 C; 411 G; 484 T; 0 U; 0 Other;

Query Match 100.0%; Score 1760; DB 12; Length 1760;

Best Local Similarity 100.0%; Fred. No. 0;

Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTCTTATTCCTTTGTAAATTCATTCATTCTTAATCTTTGCAAAAATATGTTCA 60
DB 1 CTCTCTTATTCCTTTGTAAATTCATTCATTCTTAATCTTTGCAAAAATATGTTCA 60
QY 61 TCGAGATTTTAAAGTTGAGTGTCTTAATGTGAGTACCGAGACTGATTCAGTCCG 120
DB 61 TCGAGATTTTAAAGTTGAGTGTCTTAATGTGAGTACCGAGACTGATTCAGTCCG 120
QY 121 TGTACACTACGAAACCGAATCTTTCACGAGAACGAAATGCGACCTTACGTGGA 180
DB 121 TGTACACTACGAAACCGAATCTTTCACGAGAACGAAATGCGACCTTACGTGGA 180
QY 181 TTGTCAAAACCAATCTGTCAATACGAATTTAAACCAACATCCATGTTCTTAATTG 240
DB 181 TTGTCAAAACCAATCTGTCAATACGAATTTAAACCAACATCCATGTTCTTAATTG 240
QY 241 GGGTAATGCTTGTGGGTGGGTTGGAACACGGCTCAACCTCACCGGTGGTGTATTG 300
DB 241 GGGTAATGCTTGTGGGTGGGTTGGAACACGGCTCAACCTCACCGGTGGTGTATTG 300
QY 301 CTAAACGAGAGGCAATTTATGGGCTACAAAGACAAATTCACAAAGCAATTCCTTG 360
DB 301 CTAAACGAGAGGCAATTTATGGGCTACAAAGACAAATTCACAAAGCAATTCCTTG 360
QY 361 GCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGCTCTTCCAGGGAGGAAATCTATG 420
DB 361 GCTCCCTCAACCAAGCTCAGCTATCCGAGTTGGGCTCTTCCAGGGAGGAAATCTATG 420
QY 421 CCCCATTTCAAGAGCCTGCTTCCATGTGTTACCTTGACGACATTTGTGTTGGGGATGG 480
DB 421 CCCCATTTCAAGAGCCTGCTTCCATGTGTTACCTTGACGACATTTGTGTTGGGGATGG 480
QY 481 AATATGACCAATGAACTGTGCTGATGCCATGGCCAAAGGTGTTTGACATCGATT 540
DB 481 AATATGACCAATGAACTGTGCTGATGCCATGGCCAAAGGTGTTTGACATCGATT 540
QY 541 TGCAGAGAGGTTGGGCTTACATGGAATCCATGCTTCCCGGAACTTATGACC 600
DB 541 TGCAGAGAGGTTGGGCTTACATGGAATCCATGCTTCCCGGAACTTATGACC 600
QY 601 CGAATTCATTCGTGCAACCAAGAGAGCTGTCCAACAGTCATCAAGGGCACAAAGC 660

DB 601 CGAATTCATTCGTGCAACCAAGAGAGCTGTCCAACAGTCATCAAGGGCACAAAGC 660
QY 661 AAGAGCAAGTTCAACCAATCATCAAAAGCATCAAGGCGTTTAAAGAACCCCAAAATGG 720
DB 661 AAGAGCAAGTTCAACCAATCATCAAAAGCATCAAGGCGTTTAAAGAACCCCAAAATGG 720
QY 721 ACAAGTGGTTGTACTGTGACCTGCAACAGAGAGGTACAGTAATTTGGTGTGGCC 780
DB 721 ACAAGTGGTTGTACTGTGACCTGCAACAGAGAGGTACAGTAATTTGGTGTGGCC 780
QY 781 TTAATGACACCATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAGATTTTC 840
DB 781 TTAATGACACCATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAGATTTTC 840
QY 841 CTTCACCTGTGATGCCATTTGCTGTGTATGAAAATGTTCTTCAATATGGAAGCC 900
DB 841 CTTCACCTGTGATGCCATTTGCTGTGTATGAAAATGTTCTTCAATATGGAAGCC 900
QY 901 CTCAGAACACTTTTGTATACAGGGCTGATTTGATCTTGCATGCGAGAACCTTGTATTG 960
DB 901 CTCAGAACACTTTTGTATACAGGGCTGATTTGATCTTGCATGCGAGAACCTTGTATTG 960
QY 961 GTGAGATGACTTCAAGTGTGTGACACCAAAATGAATCTGTGTGTTGATTTCTTG 1020
DB 961 GTGAGATGACTTCAAGTGTGTGACACCAAAATGAATCTGTGTGTTGATTTCTTG 1020
QY 1021 TGGGGGCTGGTATCAAGCCAACTATATGTCAGTTACATCAACATCTGGGAAACAATGATG 1080
DB 1021 TGGGGGCTGGTATCAAGCCAACTATATGTCAGTTACATCAACATCTGGGAAACAATGATG 1080
QY 1081 GTATGAATCTTTCGGCTCCACAACTTTCGGTCCAGAGAAATCTCCAAGAGCAACGTTG 1140
DB 1081 GTATGAATCTTTCGGCTCCACAACTTTCGGTCCAGAGAAATCTCCAAGAGCAACGTTG 1140
QY 1141 TTGATGATATGTTCAACGCAATGCTCTATGACCTGTGTGAAATCATCAAGCCATG 1200
DB 1141 TTGATGATATGTTCAACGCAATGCTCTATGACCTGTGTGAAATCATCAAGCCATG 1200
QY 1201 TTGATGATATGTTCAACGCAATGCTCTATGACCTGTGTGAAATCATCAAGCCATG 1260
DB 1201 TTGATGATATGTTCAACGCAATGCTCTATGACCTGTGTGAAATCATCAAGCCATG 1260
QY 1261 CTTCAAGATATTCATGAGGTGGAAGAAAGACACCATTTGTTGCACAAACATGCGAGATT 1320
DB 1261 CTTCAAGATATTCATGAGGTGGAAGAAAGACACCATTTGTTGCTCAACATGCGAGATT 1320
QY 1321 CCTCTTACGCTGCTCTATTTATCTTGACCTTGTGCTTCTTGTGAGCTCAGCACTAGAA 1380
DB 1321 CCTCTTACGCTGCTCTATTTATCTTGACCTTGTGCTTCTTGTGAGCTCAGCACTAGAA 1380
QY 1381 TCGAGTTAAAGCTGAAAATGAGGAAAATTCACATCTTCCACCCAGATTGCTACATCC 1440
DB 1381 TCGAGTTAAAGCTGAAAATGAGGAAAATTCACATCTTCCACCCAGATTGCTACATCC 1440
QY 1441 TCAGCTACCTCAACCAAGGCTCTCTGTGTTCCACCGGGTACACAGTGTGATGCAATTGT 1500
DB 1441 TCAGCTACCTCAACCAAGGCTCTCTGTGTTCCACCGGGTACACAGTGTGATGCAATTGT 1500
QY 1501 CAAGACGCTGCAATGCTGAAAAATATGAGAGGCTTGTGTGATGTGCCCCCAGAGA 1560
DB 1501 CAAGACGCTGCAATGCTGAAAAATATGAGAGGCTTGTGTGATGTGCCCCCAGAGA 1560
QY 1561 AATAAGATTTCTGAGATCAAGAGACATGAGGACGAGAAATATATATGTTGGGGTATG 1620
DB 1561 AATAAGATTTCTGAGATCAAGAGACATGAGGACGAGAAATATATATGTTGGGGTATG 1620
QY 1621 CCTAGCTGAATGTTTATGTTAATATATGTTTGTGTTAATTTTGAAGTATTTGAA 1680
DB 1621 CCTAGCTGAATGTTTATGTTAATATATGTTTGTGTTAATTTTGAAGTATTTGAA 1680
QY 1681 TGCATCACTTCATTAATGCTTTAGAGCGGGCATATTTCTGTTTATGAGAAATGAAATG 1740

DB 1681 TGCATCAGCTTCAATATGCTTAGAGCGGCGATATTCGTTTACAGGACATGATG 1740
QY 1741 AATGATGATTAATTTTGTGT 1760
DB 1741 AATGATGATTAATTTTGTGT 1760

RESULT 2
AAV62440
ID AAV62440 standard; cDNA; 1782 BP.
XX
XX AAV62440;
AC
XX 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)
XX
DE Soybean wild-type myo-inositol 1-phosphate synthase cDNA.
XX
XX Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KM phytic acid; ds.
XX
XX Glycine max; line LR13.
OS
XX
XX Key Location/Qualifiers
FT CDS 54..1586
FT /*tag= a
XX
XX W09845448-A1.
XX
XX 15-OCT-1998.
PD
XX
XX 07-APR-1998; 98WO-US006822.
XX
XX 08-APR-1997; 97US-00835751.
PR
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Hitz WD, Sebastian SA;
PI
XX
XX WPI; 1998-568353/48.
DR P-PSDB; AAW79740.
XX
XX Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
PT phytic acid, etc.
XX
XX
XX Example 5; Page 44-45; 63p; English.
PS
XX
XX This is the nucleotide sequence of cDNA encoding the wild-type soybean
CC myo-inositol 1-phosphate synthase (MI 1-PS) present in clone pSbm-1ps
CC (ATCC 97970). The clone was isolated from a cDNA library of soybean line
CC LR13 by hybridisation to a probe made from MI 1-PS cDNA of Arabidopsis
CC thaliana. MI 1-PS is involved in glucose metabolism to phytic acid,
CC raffinose and stachyose. A mutant MI 1-PS nucleic acid (see AAV62443) has
CC been identified in soybean line LR13, a mutagenised line of low raffinose
CC saccharide phenotype. Sequencing revealed a single base change mutation
CC (G to T at base 1241) in the LR13 sequence. The mutation results in a
CC seed phenotype of very low raffinose saccharide sugars, very high sucrose
CC and low phytic acid. The nucleic acid is used to alter the raffinose
CC saccharide, sucrose, phytic acid and inorganic phosphate content of
CC soybean seeds, leading to useful soybean products, e.g. a seed phytic
CC acid content of less than 17 ug/g, a seed content of raffinose and
CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
CC greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)
CC
XX
SQ Sequence 1782 BP; 516 A; 371 C; 411 G; 484 T; 0 U; 0 Other;

Query Match 100.0%; Score 1760; DB 2; Length 1782;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTCTTATATCCTTTGTAATTCATTCATTCCTTAATCTTTGTAATAATTAATGTCA 60
|||||

DB 1 CTTCTTATATCCTTTGTAATTCATTCATTCCTTAATCTTTGTAATAATTAATGTCA 60
QY 61 TCGAATATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCCGAGATCGATTCAATCCG 120
|||||
DB 61 TCGAATATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCCGAGATCGATTCAATCCG 120
QY 121 TGTACAACTACGAAACCAACGAACTTGTACAGAGAACGAAATGACCTATCACTGGA 180
|||||
DB 121 TGTACAACTACGAAACCAACGAACTTGTGTACAGAGAACGAAATGACCTATCACTGGA 180
QY 181 TTGTCAAAACCAATCTGTCAATAAGATTAAACCAATCCATGTTCTTAATTAG 240
|||||
DB 181 TTGTCAAAACCAATCTGTCAATAAGATTAAACCAATCCATGTTCTTAATTAG 240
QY 241 GGGTAAATGCTTGGGTTGGGTTGGAAACAAGGCTCAACCTCACCGGTGTTATTG 300
|||||
DB 241 GGGTAAATGCTTGGGTTGGGTTGGAAACAAGGCTCAACCTCACCGGTGTTATTG 300
QY 301 CTAAACGAGAGGCAATTCATGGGGCTACAAAGGACAAAGATTCAACAAAGCAATTA 360
|||||
DB 301 CTAAACGAGAGGCAATTCATGGGGCTACAAAGGACAAAGATTCAACAAAGCAATTA 360
QY 361 GCTCCCTCAACCAAGGCTCAGCTATCCAGTTGGGTCTTCCAGAGGAGAAATCTATG 420
|||||
DB 361 GCTCCCTCAACCAAGGCTCAGCTATCCAGTTGGGTCTTCCAGAGGAGAAATCTATG 420
QY 421 CCCCATTCAAAGACCTGCTTCCAAATGTTAACCTTGACGACATGTTGTTGGGGATGG 480
|||||
DB 421 CCCCATTCAAAGACCTGCTTCCAAATGTTAACCTTGACGACATGTTGTTGGGGATGG 480
QY 481 AATATCAACAATGAACCTGGGTGATGCAATGGCCAGAGGCAAAAGGTTTGAATGAT 540
|||||
DB 481 AATATCAACAATGAACCTGGGTGATGCAATGGCCAGAGGCAAAAGGTTTGAATGAT 540
QY 541 TCGAAGAGAGTTGAGGCTTACATGGAATCCATGCTTCCATCCCGGAATCTATGAC 600
|||||
DB 541 TCGAAGAGAGTTGAGGCTTACATGGAATCCATGCTTCCATCCCGGAATCTATGAC 600
QY 601 CGGATTCATTGCTGCGAACCAAGAGAGGCTGCCAACAGCTCATCAAGGCAACAGC 660
|||||
DB 601 CGGATTCATTGCTGCGAACCAAGAGAGGCTGCCAACAGCTCATCAAGGCAACAGC 660
QY 661 AAGAGCAAGTTCAACAAATCATCAAAAGCATCAAGGCGTTTAAAGAAAGCACAAGTGG 720
|||||
DB 661 AAGAGCAAGTTCAACAAATCATCAAAAGCATCAAGGCGTTTAAAGAAAGCACAAGTGG 720
QY 721 ACAAGGTGTTGACTGTGAGCTGCAACACAGAGAGTACGTAATTTGTTGGGGCC 780
|||||
DB 721 ACAAGGTGTTGACTGTGAGCTGCAACACAGAGAGTACGTAATTTGTTGGGGCC 780
QY 781 TTAATGACACATGAGAAATCTCTTGGCTGTGTGACAGAAATGAGCTGAGATTTCTC 840
|||||
DB 781 TTAATGACACATGAGAAATCTCTTGGCTGTGTGACAGAAATGAGCTGAGATTTCTC 840
QY 841 CTTCAACCTTGATACCATGCTTGTGTGTAAGAAATTTCTTCAATTAATGAAAGCC 900
|||||
DB 841 CTTCAACCTTGATACCATGCTTGTGTGTAAGAAATTTCTTCAATTAATGAAAGCC 900
QY 901 CTCAGAACATTTTGTACAGAGGCTGATGATCTTGCCATCCGAGAGAACATCTGATTG 960
|||||
DB 901 CTCAGAACATTTTGTACAGAGGCTGATGATCTTGCCATCCGAGAGAACATCTGATTG 960
QY 961 GTGAGATGACTTCAAGAGTGTCAAGACCAAAATGAATCTGTGTTGTTGATTTCTTG 1020
|||||
DB 961 GTGAGATGACTTCAAGAGTGTCAAGACCAAAATGAATCTGTGTTGTTGATTTCTTG 1020
QY 1021 TGGGGGCTGTATCAAGCCAAATCTATATGATGATTAACAACATCTGGAACCAATGATG 1080
|||||
DB 1021 TGGGGGCTGTATCAAGCCAAATCTATATGATGATTAACAACATCTGGAACCAATGATG 1080
QY 1081 GTATGAATCTTCGGGCTCAACAACTTTCGTTCCAGAGAAATCTCAAGAGCAACGTTG 1140
|||||
DB 1081 GTATGAATCTTCGGGCTCAACAACTTTCGTTCCAGAGAAATCTCAAGAGCAACGTTG 1140


```

Qy 414 ATCTATGCCCATCTCAAGAGCGCTTCCAAATGGTTAACCTGACGACATTTGTTGGG 473
Db 361 ATCTATGCCCATCTCAAGAGCGCTTCCAAATGGTTAACCTGACGACATTTGTTGGG 420
Qy 474 GGAATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGAGCAAGTGTGGAC 533
Db 421 GGAATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGAGCAAGTGTGGAC 480
Qy 534 ATGATTTTGCAGAGCAAGTGGAGGCTTACATGGAATCCATGCTTCCACTCCCCGAATC 593
Db 481 ATGATTTTGCAGAGCAAGTGGAGGCTTACATGGAATCCATGCTTCCACTCCCCGAATC 540
Qy 594 TATGACCCGGATTTCAATGCTGCGCAACAGAGAGGCTGCCAACAACGTCATCAAGGCG 653
Db 541 TATGACCCGGATTTCAATGCTGCGCAACAGAGAGGCTGCCAACAACGTCATCAAGGCG 600
Qy 654 ACAAGCAAGAGCAAGTTCACAAATCATCAAGACATCAAGGCGTTTAAAGAAAGCACCC 713
Db 601 ACAAGCAAGAGCAAGTTCACAAATCATCAAGACATCAAGGCGTTTAAAGAAAGCACCC 660
Qy 714 AAAGTGAACAAGGTGGTGTGACTGTGCAACAGAGAGGTACAGTAATTTGGTT 773
Db 661 AAAGTGAACAAGGTGGTGTGACTGTGCAACAGAGAGGTACAGTAATTTGGTT 720
Qy 774 GTGGGCGCTTAATGACCAATGGAATCTTGGCTGTGGACAAGAAATGAGCGTGAG 833
Db 721 GTGGGCGCTTAATGACCAATGGAATCTTGGCTGTGGACAAGAAATGAGCGTGAG 780
Qy 834 ATTTCTCTCTCCACCTTGTATGCGCATTTGTTATGGAATAATTTCTTTCATTAAAT 893
Db 781 ATTTCTCTCTCCACCTTGTATGCGCATTTGTTATGGAATAATTTCTTTCATTAAAT 840
Qy 894 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTTGATCTTCCATCCGAGAGAACCT 953
Db 841 GGAAGCCCTCAGAACCTTTTGTACCAAGGCTGATTTGATCTTCCATCCGAGAGAACCT 900
Qy 954 TTGATTTGGGAGATGACTTCAAGAGTGGTCAGACCAAAATGAAATCTGTTGGTGAAT 1013
Db 901 TTGATTTGGGAGATGACTTCAAGAGTGGTCAGACCAAAATGAAATCTGTTGGTGAAT 960
Qy 1014 TTCTTTGTGGGGCTGTATCAAGCCAACTATATGTCATTAACAACATCTGGGAAAC 1073
Db 961 TTCTTTGTGGGGCTGTATCAAGCCAACTATATGTCATTAACAACATCTGGGAAAC 1020
Qy 1074 AATGATGTAATGTAATTTTGGGCTCCACAAACTTTCCTTCCAGGAAATTTCCAAAGAC 1133
Db 1021 AATGATGTAATGTAATTTTGGGCTCCACAAACTTTCCTTCCAGGAAATTTCCAAAGAC 1080
Qy 1134 AACGTTTGTAGATATGTCACACAGCANTGCCATCTTATGAGCTGTGAACATCCA 1193
Db 1081 AACGTTTGTAGATATGTCACACAGCANTGCCATCTTATGAGCTGTGAACATCCA 1140
Qy 1194 GACCATGTTGTTTATTAAGTATGTCCTTACGTAGGGGACAGCAAGAGAGCCATGAT 1253
Db 1141 GACCATGTTGTTTATTAAGTATGTCCTTACGTAGGGGACAGCAAGAGAGCCATGAT 1200
Qy 1254 GAGTACACTTCAGAGATATTCATGGGTGAAGAGACCAATGTTTGGACAACACATCC 1313
Db 1201 GAGTACACTTCAGAGATATTCATGGGTGAAGAGACCAATGTTTGGACAACACATCC 1260
Qy 1314 GAGGATTCCTCTGAGCTGCTCTATTAATCTTGGATTTGCTCTTCTGAGCTCAGC 1373
Db 1261 GAGGATTCCTCTGAGCTGCTCTATTAATCTTGGATTTGCTCTTCTGAGCTCAGC 1320
Qy 1374 ACTAGATTCAGATTTAAAGCTGAAATGAGGAAATTCACATCTTCCACCAAGTTGCT 1433
Db 1321 ACTAGATTCAGATTTAAAGCTGAAATGAGGAAATTCACATCTTCCACCAAGTTGCT 1380
Qy 1434 ACCATCTCAGCTACTCTCAACCAAGGCTCTCTGTTTCCACCGGGTACACCAAGTGTGAAT 1493
Db 1381 ACCATCTCAGCTACTCTCAACCAAGGCTCTCTGTTTCCACCGGGTACACCAAGTGTGAAT 1440
Qy 1494 GCATTGTCAAAGCAGGTCGCAATGCTGGAATAATAGGGCTTGTGTGATTGGCC 1553

```

```

Db 1441 GCATTGTCAAAGCAGGTCGCAATGCTGGAATAATAGGGCTTGTGTGATTGGCC 1500
Qy 1554 CCAGAGATTAACATGATTTCTCGAGTACAAGTGA 1586
Db 1501 CCAGAGATTAACATGATTTCTCGAGTACAAGTGA 1533

RESULT 4
ADS81999
ID ADS81999 standard; cDNA, 1533 BP.
XX
XX
AC ADS81999;
XX
XX
DT 18-NOV-2004 (first entry)
XX
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 3.
XX
XX
KM Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phylic acid;
XX
XX
OS raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
XX
OS Glycine max; line 29004JP01.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag=a
FT /product="myo-inositol 1-phosphate synthase"
XX
XX
PN US2003074685-A1.
XX
XX
PD 17-APR-2003.
XX
XX
PF 11-MAR-2002; 2002US-00025003.
XX
XX
PR 08-APR-1997; 97US-00835751.
XX
XX
PR 07-APR-1998; 98WO-US006822.
XX
XX
PA (HITZ/) HITZ W D.
XX
XX
PA (SEBA/) SEBASTIAN S A.
XX
XX
PI Hitz WD, Sebastian SA;
XX
XX
DR WPI; 2004-639957/62.
XX
XX
DR P-PSDB; ADS82000.
XX
XX
PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
XX
XX
PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
XX
XX
PT phylic acid and inorganic phosphate content of soybean seeds.
XX
XX
PS Example 8; SEQ ID NO 9; 34bp; English.
XX
XX
CC The invention relates to an isolated nucleic acid fragment encoding a
XX
XX
CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
XX
XX
CC phosphate synthase having decreasing capacity for the synthesis for myo-
XX
XX
CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
XX
XX
CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
XX
XX
CC complement, subfragment or the complement of the subfragment, operably
XX
XX
CC linked to suitable regulatory sequences, where expression of the chimeric
XX
XX
CC gene results in a decrease in expression of an endogenous or native gene
XX
XX
CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
XX
XX
CC comprising the chimeric gene (with a heritable phenotype of a seed phylic
XX
XX
CC acid content of less than 17 micromol/g, a seed content of raffinose plus
XX
XX
CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
XX
XX
CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
XX
XX
CC from the plant, making a soybean plant with the heritable phenotype
XX
XX
CC (comprising crossing LR33 or the plant comprising the chimeric gene with
XX
XX
CC an elite soybean plant and selecting a progeny plant of the cross of
XX
XX
CC crossing step that has a heritable phenotype as mentioned above), seeds
XX
XX
CC of soybean plant made by the above method, a soy protein product derived
XX
XX
CC from seeds of a soybean plant (homozygous for one or more gene encoding a
XX
XX
CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
XX
XX
CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
XX
XX
CC heritable phenotype as mentioned above), and making or producing a

```

CC soybean protein product derived from seeds of a soybean plant with
CC heritable phenotype as mentioned above. The nucleic acid is useful for
CC altering raffinose saccharide, sucrose, phytic acid and inorganic
CC phosphate content of soybean seeds thus leading to valuable and useful
CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a wild-type myo-
CC inositol 1-phosphate synthase.

XX Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Query Match 87.1%; Score 1533; DB 13; Length 1533;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 54 ATGTTTCATCGAGAAATTTTAAAGTTGAGTCTCTTAATGTGAATACACCGAGCTAGAT 113
DB 1 ATGTTTCATCGAGAAATTTTAAAGTTGAGTCTCTTAATGTGAATACACCGAGCTAGAT 60
QY 114 CAGTCCGTTGACAACTACGAAACCAACCGAAGTGTTCAGAGAACAGAAATGSCACTAT 173
DB 61 CAGTCCGTTGACAACTACGAAACCAACCGAAGTGTTCAGAGAACAGAAATGSCACTAT 120
QY 174 CAGTGGATTGTCAAAACCCAAATCTGTCAAAATAGAAATTTAAACCAACATCATGTTCT 233
DB 121 CAGTGGATTGTCAAAACCCAAATCTGTCAAAATAGAAATTTAAACCAACATCATGTTCT 180
QY 234 AAATTAGGGGTAAATGTTGGGTGGGGTGGAGAAACAAGGCTCAACCTTCACCGGTGT 293
DB 181 AAATTAGGGGTAAATGTTGGGTGGGGTGGAGAAACAAGGCTCAACCTTCACCGGTGT 240
QY 294 GTTATTGCTAACCGAGAGGCAATTTGATGGGCTACAAAGACAAAGATTCAACAGGCAT 353
DB 241 GTTATTGCTAACCGAGAGGCAATTTGATGGGCTACAAAGACAAAGATTCAACAGGCAT 300
QY 354 TACTTTGGCTCCCTCAACCCAGGCTCAGTATCCGAGTTGGGTCTTCCAGGAGAGAA 413
DB 301 TACTTTGGCTCCCTCAACCCAGGCTCAGTATCCGAGTTGGGTCTTCCAGGAGAGAA 360
QY 414 ATCTATGCCCCCATTCAGAGGCTGCTTCCAAATGTTTAAACCTTCAGACACTTGTGTTGG 473
DB 361 ATCTATGCCCCCATTCAGAGGCTGCTTCCAAATGTTTAAACCTTCAGACACTTGTGTTGG 420
QY 474 GGATGGGATATACGAACAATGAACCTGGCTGATGCCATGGCCAGGCAAAAGGTTTGAC 533
DB 421 GGATGGGATATACGAACAATGAACCTGGCTGATGCCATGGCCAGGCAAAAGGTTTGAC 480
QY 534 ATCGATTTCAGAGAGGATTGAGGCTTTACATGAATCCATGCTTCACCTCCCGGAATC 593
DB 481 ATCGATTTCAGAGAGGATTGAGGCTTTACATGAATCCATGCTTCACCTCCCGGAATC 540
QY 594 TATGACCCCGGATTTCTTCTGTCCCAACAGAGACCGTCCAAACAACCTCATCAAGGGC 653
DB 541 TATGACCCCGGATTTCTTCTGTCCCAACAGAGACCGTCCAAACAACCTCATCAAGGGC 600
QY 654 ACAAGCAAGAGCAATTCACAAATCATCAAGACATCAAGGCGTTTAAAGAGGACAC 713
DB 601 ACAAGCAAGAGCAATTCACAAATCATCAAGACATCAAGGCGTTTAAAGAGGACAC 660
QY 714 AAAGTGACAAGGTGTTGACTGTGACTGCAACACAGAGAGTCAAGTAATTTGGTT 773
DB 661 AAAGTGACAAGGTGTTGACTGTGACTGCAACACAGAGAGTCAAGTAATTTGGTT 720
QY 774 GTGGGCTTAAATACCAATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAG 833
DB 721 GTGGGCTTAAATACCAATGAGAAATCTTTGGCTGTGTGACAGAAATGAGGCTGAG 780
QY 834 ATTTCTCTTCCACTTGTATGSCATTGCTGTGTATGGAATAATGTTCTTTCAATTAAT 893
DB 781 ATTTCTCTTCCACTTGTATGSCATTGCTGTGTATGGAATAATGTTCTTTCAATTAAT 840
QY 894 GGAAGCCTCAGAAACCTTTTGTACAGAGGCTGATGATCTTTGCAATCCGAGGAACACT 953
DB 894 GGAAGCCTCAGAAACCTTTTGTACAGAGGCTGATGATCTTTGCAATCCGAGGAACACT 953
```

```
DB 841 GGAAGCCTCAGAAACCTTTTGTATCAGAGGCTGATGATCTTTGCCATCCAGAGAAACACT 900
QY 954 TTGATTGGTGGAGATGACTTCAAGAGTGGTCAGACCAAAATGAATCTGTGTGGTTGAT 1013
DB 901 TTGATTGGTGGAGATGACTTCAAGAGTGGTCAGACCAAAATGAATCTGTGTGGTTGAT 960
QY 1014 TTCTTGTGGGGGCTGTATCAAGCAACATCTATAGTCAATTACACCATCTGGGAAAC 1073
DB 961 TTCTTGTGGGGGCTGTATCAAGCAACATCTATAGTCAATTACACCATCTGGGAAAC 1020
QY 1074 AATGATGATTAATCTTTGGGCTCCAAACCTTTCCGTTCCAGGAATCTCCAGAGG 1133
DB 1021 AATGATGATTAATCTTTGGGCTCCAAACCTTTCCGTTCCAGGAATCTCCAGAGG 1080
QY 1134 AACGTTTGTATGATATGATGTCACAGCAATGCCATCTCTATAGAGCTGTGTAACATCA 1193
DB 1081 AACGTTTGTATGATATGATGTCACAGCAATGCCATCTCTATAGAGCTGTGTAACATCA 1140
QY 1194 GACCATGTTGTTGTTATTAATGATGTGCTTACGTAAGGAGACAGCAAGAGCCATGGAT 1253
DB 1141 GACCATGTTGTTGTTATTAATGATGTGCTTACGTAAGGAGACAGCAAGAGCCATGGAT 1200
QY 1254 GAGTACACTTCAGAGATATTCATAGGTGGAAGAGCAACATTTGTTGCAACACATGAC 1313
DB 1201 GAGTACACTTCAGAGATATTCATAGGTGGAAGAGCAACATTTGTTGCAACACATGAC 1260
QY 1314 GAGATTCCTCTTGTAGCTGCTCTTATATCTTGAAGTGTCTTCTTGTGAGCTCAGC 1373
DB 1261 GAGATTCCTCTTGTAGCTGCTCTTATATCTTGAAGTGTGTCTTGTGAGCTCAGC 1320
QY 1374 ACTAGAAATCGAGTTTAAAGCTGAAATAGAGGAAATTCACATCTCCACCAGTGTCT 1433
DB 1321 ACTAGAAATCGAGTTTAAAGCTGAAATAGAGGAAATTCACATCTCCACCAGTGTCT 1380
QY 1434 ACCATCTCTAGTACTTACCAACAGGCTCTCTGTGTTCCACCGGGATCACCGTGTGAT 1493
DB 1381 ACCATCTCTAGTACTTACCAACAGGCTCTCTGTGTTCCACCGGGATCACCGTGTGAT 1440
QY 1494 GCATTGTCAAAGCAGCTGCAATGCTGGAATAACATATAGAGGCTTGTGTGATTTGGCC 1553
DB 1441 GCATTGTCAAAGCAGCTGCAATGCTGGAATAACATATAGAGGCTTGTGTGATTTGGCC 1500
QY 1554 CCAGAGAAATTAACATGATTTCTGAGTACAAATGA 1586
DB 1501 CCAGAGAAATTAACATGATTTCTGAGTACAAATGA 1533
RESULT 5
ADS81993
ID ADS81993 standard; cDNA, 1533 BP.
XX
AC ADS81993;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 1.
XX
KW Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
XX
OS Glycine max; cultivar Wye.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
XX
PN US2003074685-A1.
XX
PD 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
```

PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98MO-US006822.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 PI
 PI Hitz WD, Sebastian SA;
 XX
 DR MPI, 2004-639957/62.
 DR P-PSDB; ADS81994.

Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.

Claim 2; SEQ ID NO 1; 34bp; English.

The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR3), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR3 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homologous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with a heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful soybean products, since the presence of high concentration of raffinose oligosaccharides in soy plants (and other legumes) can lead to flatulence when consumed by humans. The present sequence encodes a wild-type myo-inositol 1-phosphate synthase.

Sequence 1533 BP; 429 A; 344 C; 368 G; 392 T; 0 U; 0 Other;

Query Match 87.1%; Score 1533; DB 13; Length 1533;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 AGCTTCATGAGAAATTTTAAAGTTGAGTGTCTTAATGTAAGTACCGAGCTGAGATT 113
 DB 1 ATGTTTCATGAGAAATTTTAAAGTTGAGTGTCTTAATGTAAGTACCGAGCTGAGATT 60
 QY 114 CAGTCCGTGTACCACTACGAAACCAACGAACTGTTCAGAGAAACGAGATGSCACTAT 173
 DB 61 CAGTCCGTGTACCACTACGAAACCAACGAACTGTTCAGAGAAACGAGATGSCACTAT 120
 QY 174 CAGTGGATGTCAAAACCAATCTGTCAATACGAATTTTAAACCAACATTCATGTTCTT 233
 DB 121 CAGTGGATGTCAAAACCAATCTGTCAATACGAATTTTAAACCAACATTCATGTTCTT 180
 QY 234 AAATTAGGGGTATATGTTGGTGGTGGTGAAGAAAGCGCTCAACCTCACCGGTGGT 293
 DB 181 AAATTAGGGGTATATGTTGGTGGTGGTGAAGAAAGCGCTCAACCTCACCGGTGGT 240
 QY 294 GTTATTGCTAACGAGAGGAGCTATTCATGGGCTACAAAGACAAAGTTCAACAGCAAT 353
 DB 241 GTTATTGCTAACGAGAGGAGCTATTCATGGGCTACAAAGACAAAGTTCAACAGCAAT 300

QY 354 TACTTTGGCTCCCTGACCCCAAGCCTCAGCTATCCGAGTTGGGCTTCCAGGGAGAGGAA 413
 DB 301 TACTTTGGCTCCCTGACCCCAAGCCTCAGCTATCCGAGTTGGGCTTCCAGGGAGAGGAA 360
 QY 414 ATCTATGCCCATTTCAAGAGCCTGCTTCAATGTTTAAACCTGACGACATTTGTTGGG 473
 DB 361 ATCTATGCCCATTTCAAGAGCCTGCTTCAATGTTTAAACCTGACGACATTTGTTGGG 420
 QY 474 GGAATGGATATACGACAACTGAAACCTGCTGATGCCATGGCCAGAGGCAAGGTTTGAAC 533
 DB 421 GGATGGATATACGACAACTGAAACCTGCTGATGCCATGGCCAGAGGCAAGGTTTGAAC 480
 QY 534 ATCGATTTGCAAGACAGTGGAGCCTTACATGAAATCCATGCTTCCATCCCCCGAATC 593
 DB 481 ATCGATTTGCAAGACAGTGGAGCCTTACATGAAATCCATGCTTCCATCCCCCGAATC 540
 QY 594 TATGACCCGGATTTTCATTTGCTGCCAAACGAGAGCGTGCCAAACGTCATCAAGGCG 653
 DB 541 TATGACCCGGATTTTCATTTGCTGCCAAACGAGAGCGTGCCAAACGTCATCAAGGCG 600
 QY 654 ACAAAAGCAAGACCAAGTTCAACAAATCATCAAAAGACATCAAGCGCTTAAAGAGCCACC 713
 DB 601 ACAAAAGCAAGACCAAGTTCAACAAATCATCAAAAGACATCAAGCGCTTAAAGAGCCACC 660
 QY 714 AAAGTGAACAGGTGTTTACTGTGACCTGCCAACACAGAGGTACGTAATTTGGTT 773
 DB 661 AAAGTGAACAGGTGTTTACTGTGACCTGCCAACACAGAGGTACGTAATTTGGTT 720
 QY 774 GTGGGCTTATGACACCATGAGAAATCTTGGCTGCTGTGAGCAAGAAATGAGGCTGAG 833
 DB 721 GTGGGCTTATGACACCATGAGAAATCTTGGCTGCTGTGAGCAAGAAATGAGGCTGAG 780
 QY 834 ATTTCTCTTCCACCTTGTATGCAATGCTTGTATGAAATGTTCTTTCAATTAAT 893
 DB 781 ATTTCTCTTCCACCTTGTATGCAATGCTTGTATGAAATGTTCTTTCAATTAAT 840
 QY 894 GGAAAGCCCTCAAGAACTTTTGTACAGAGGCTGATGATCTTGGCATCGGAGAACT 953
 DB 841 GGAAAGCCCTCAAGAACTTTTGTACAGAGGCTGATGATCTTGGCATCGGAGAACT 900
 QY 954 TTGATTTGGTGAATGACTTCAAGAGTGGTGCAGACCAAAATGAAATCTGTTGGTGAAT 1013
 DB 901 TTGATTTGGTGAATGACTTCAAGAGTGGTGCAGACCAAAATGAAATCTGTTGGTGAAT 960
 QY 1014 TTCTTTGGGGGGCTGTATCAAGCCAACTATATGTCATGTAACCAATCTGGGAAAC 1073
 DB 961 TTCTTTGGGGGGCTGTATCAAGCCAACTATATGTCATGTAACCAATCTGGGAAAC 1020
 QY 1074 AATGATGATGATATCTTTGGCTCCAAACTTTTCCGTTCCAGGAAATCTCCAGAGC 1133
 DB 1021 AATGATGATGATATCTTTGGCTCCAAACTTTTCCGTTCCAGGAAATCTCCAGAGC 1080
 QY 1134 AAGCTTTGATGATATGTTGTCAACGAAATGCAATCTTATAGAGCTGGTGAATCA 1193
 DB 1081 AAGCTTTGATGATATGTTGTCAACGAAATGCAATCTTATAGAGCTGGTGAATCA 1140
 QY 1194 GACCAATGTTGTTATTAATGATGTGCTTACGTAGGGGACAGAGAGACCAATGAT 1253
 DB 1141 GACCAATGTTGTTATTAATGATGTGCTTACGTAGGGGACAGAGAGACCAATGAT 1200
 QY 1254 GAGTACACTTCAAGATATTCATGAGTGAAGAGACCAATGTTTGGCAACACATGC 1313
 DB 1201 GAGTACACTTCAAGATATTCATGAGTGAAGAGACCAATGTTTGGCAACACATGC 1260
 QY 1314 GAGATATCCCTCTTACGCTGCTCTATTAATCTTGAACCTTGATCTTGTAGCTCAGC 1373
 DB 1261 GAGATATCCCTCTTACGCTGCTCTATTAATCTTGAACCTTGATCTTGTAGCTCAGC 1320
 QY 1374 ACTAGATCGAATTTAAAGCTGAAATGAGGAAATTCATCTCAATTCACCACTAGTGT 1433
 DB 1321 ACTAGATCGAATTTAAAGCTGAAATGAGGAAATTCATCTCAATTCACCACTAGTGT 1380

QY 1434 ACCATCTCAGCTACCTCAACGAAGGCTCCTCTGTGTTCCACCGGGTACACCACTGGTGAAT 1493
| | | | |
Db 1381 ACCATCTCAGCTACCTCAACGAAGGCTCCTCTGTGTTCCACCGGGTACACCACTGGTGAAT 1440
| | | | |
QY 1494 GCATTGTCAAAGCAGCGTGCATTCGTGAAAACATATAGAGGCTTGTGTGATTTGGCC 1553
| | | | |
Db 1441 GCATTGTCAAAGCAGCGTGCATTCGTGAAAACATATAGAGGCTTGTGTGATTTGGCC 1500
| | | | |
QY 1554 CCAGAGAAATACATGATTTCTCGAGTACCAAGTGA 1586
| | | | |
Db 1501 CCAGAGAAATACATGATTTCTCGAGTACCAAGTGA 1533
| | | | |

RESULT 6
AAV62443

ID AAV62443 standard; cDNA; 1533 BP.

XX AAV62443;

DT 17-OCT-2003 (revised)
DT 02-FEB-1999 (first entry)

DE Soybean mutant myo-inositol 1-phosphate synthase cDNA.

XX Soybean; myo-inositol 1-phosphate synthase; raffinose; stachyose;
KM phytic acid; de.

XX Glycine max; line LR33.

XX WO9845448-A1.

XX 15-OCT-1998.

XX 07-APR-1998: 98WO-US006822.

XX 08-APR-1997: 97US-00835751.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

XX Hitz WD, Sebastian SA;

XX WPI; 1998-568353/48.

XX P-PSDB; AAM79741.

PT Soybean plants containing altered myo-inositol-1-phosphate gene - useful
PT for generating plants with altered levels of e.g. raffinose, stachyose,
PT phytic acid, etc.

XX Example 5; Page 48-49; 63pp; English.

CC This is the nucleotide sequence of cDNA encoding a mutant soybean myo-
CC inositol 1-phosphate synthase (MI 1-PS) present in clone LR33-10. This
CC clone was isolated from a cDNA library of soybean line LR33 by PCR
CC amplification (see AAV62441-42). Line LR33 was obtained by chemical
CC mutagenesis of wild-type soybean genome and as a reduced raffinose
CC saccharide phenotype. MI 1-PS is involved in glucose metabolism to phytic
CC acid, raffinose and stachyose. Sequencing revealed a single base change
CC mutation (G to T at base 1241) in the LR33 sequence when compared to the
CC wild-type sequence (see AAV62440). The mutation results in a seed
CC phenotype of very low raffinose saccharide sugars, very high sucrose and
CC low phytic acid. The mutated MI 1-PS nucleic acid is used to alter the
CC raffinose saccharide, sucrose, phytic acid and inorganic phosphate
CC content of soybean seeds, leading to useful soybean products, e.g. a seed
CC phytic acid content of less than 17 ug/g, a seed content of raffinose and
CC stachyose combined of less than 14.5 ug/g, and a seed sucrose content
CC greater than 200 ug/g. (Updated on 17-OCT-2003 to standardise OS field)

CC Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 87.0%; Score 1531.4; DB 2; Length 1533;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 54 ATGTTGATCGAAGATTTTAAAGTTGAGTGCTCTTAATGTGAAGTACACCGAAGCTGAGATT 113
| | | | |
Db 1 ATGTTGATCGAAGATTTTAAAGTTGAGTGCTCTTAATGTGAAGTACACCGAAGCTGAGATT 60
| | | | |
QY 114 CAGTCCGTTGTCACACTAGAAAACCGAAGCTGTTGTCACGGAACAGGAATGGACCTAT 173
| | | | |
Db 61 CAGTCCGTTGTCACACTAGAAAACCGAAGCTGTTGTCACGGAACAGGAATGGACCTAT 120
| | | | |
QY 174 CAGTGGATTGTCAAACCCCAAAATCTGTCAAAATACGAATTTAAACCAATCCATGTTCT 233
| | | | |
Db 121 CAGTGGATTGTCAAACCCCAAAATCTGTCAAAATACGAATTTAAACCAATCCATGTTCT 180
| | | | |
QY 234 AAATTAGGGGTAAATGCTTGTGGGTGGGTTGGGTAACCAACGCTCAACCTCAACGGTGT 293
| | | | |
Db 181 AAATTAGGGGTAAATGCTTGTGGGTGGGTTGGGTAACCAACGCTCAACCTCAACGGTGT 240
| | | | |
QY 294 GTTATTTGCTAACCGAAGGCAATTTCAATGGGCTACAAAGGACAGATTCACCAAGCAAT 353
| | | | |
Db 241 GTTATTTGCTAACCGAAGGCAATTTCAATGGGCTACAAAGGACAGATTCACCAAGCAAT 300
| | | | |
QY 354 TACTTTGGCTCCCTCACCACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGGAA 413
| | | | |
Db 301 TACTTTGGCTCCCTCACCACCAAGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGGAA 360
| | | | |
QY 414 ATCTATGCCCATTTCAAGAGCTGCTTCCATGCTTAACCTGACGACATGTGTGGG 473
| | | | |
Db 361 ATCTATGCCCATTTCAAGAGCTGCTTCCATGCTTAACCTGACGACATGTGTGGG 420
| | | | |
QY 474 GGATGGGATATCAGAAATGAAACCTGGCTGATGTCATGGCCAGGGCAAGGTGTTTAC 533
| | | | |
Db 421 GGATGGGATATCAGAAATGAAACCTGGCTGATGTCATGGCCAGGGCAAGGTGTTTAC 480
| | | | |
QY 534 ATCGATTTGACAGAAACAGTTGAGGCTTACATGAAATCAATGCTTCCACTCCCGGAATC 593
| | | | |
Db 481 ATCGATTTGACAGAAACAGTTGAGGCTTACATGAAATCAATGCTTCCACTCCCGGAATC 540
| | | | |
QY 594 TATGACCCGGATTTTATTGCTGCTGCCAACCAAGAGAGCTGCCACACAGCTATCAAGGC 653
| | | | |
Db 541 TATGACCCGGATTTTATTGCTGCTGCCAACCAAGAGAGCTGCCACACAGCTATCAAGGC 600
| | | | |
QY 654 ACNAAAGCAAGACAGTTCAACAAATCATCAAGACATCAAGAGCGTTAAAGAACCCACC 713
| | | | |
Db 601 ACNAAAGCAAGACAGTTCAACAAATCATCAAGACATCAAGAGCGTTAAAGAACCCACC 660
| | | | |
QY 714 AAAGTGCACAAAGTGGTGTGACTGTGACCTGCCCAACAGAGAGGTACAGTAATTTGGTT 773
| | | | |
Db 661 AAAGTGCACAAAGTGGTGTGACTGTGACCTGCCCAACAGAGAGGTACAGTAATTTGGTT 720
| | | | |
QY 774 GTGGGCTTTAATGACACCATGAGAGATCTTGTGCTGTGTGACAGAAATGAGGCTGAG 833
| | | | |
Db 721 GTGGGCTTTAATGACACCATGAGAGATCTTGTGCTGTGTGACAGAAATGAGGCTGAG 780
| | | | |
QY 834 ATTTCTCTTCCACTTGTGATGCCATTGTGTGTGTTAAGAAAAGTTCTTTCATTAAT 893
| | | | |
Db 781 ATTTCTCTTCCACTTGTGATGCCATTGTGTGTGTTAAGAAAAGTTCTTTCATTAAT 840
| | | | |
QY 894 GGAGAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTGCCATGCGAGAACACT 953
| | | | |
Db 841 GGAGAGCCCTCAGAACACTTTTGTACAGAGGCTGATGATCTTGCCATGCGAGAACACT 900
| | | | |
QY 954 TTGATTTGTGAGATGACTTCAAGAGTGCTGACACCAAAATGAAATCTGTGTTGGTAT 1013
| | | | |
Db 901 TTGATTTGTGAGATGACTTCAAGAGTGCTGACACCAAAATGAAATCTGTGTTGGTAT 960
| | | | |
QY 1014 TTCTCTTGTGGGGCTGTGATACAGCAACATTAATGTCAGTTAAACCATTTGGGAAC 1073
| | | | |
Db 961 TTCTCTTGTGGGGCTGTGATACAGCAACATTAATGTCAGTTAAACCATTTGGGAAC 1020
| | | | |
QY 1074 AATGATGATGATGATCTTTGGGCTCCACAAATCTTCCGTTCCAAAGAAATCTCAAGAGC 1133
| | | | |
Db 1021 AATGATGATGATGATCTTTGGGCTCCACAAATCTTCCGTTCCAAAGAAATCTCAAGAGC 1080
| | | | |
QY 1134 AACGTTGTGATGATATGATGATCAACAGCAATGCCATCTCTATGAGCCTGTGTAACATCCA 1193
| | | | |

```
Db 1081 AACGTTGATGATATGTCACACAGCAATGCCATCTCTATAGAGCTGTGTAACATCCA 1140
Qy 1194 GACCATGTTGTTGTTATTTAAGTATGTCCTTACGTATGGGAGACAGAAAGCCATGAT 1253
Db 1141 GACCATGTTGTTGTTATTTAAGTATGTCCTTACGTATGGGAGACAGCAATAGCCATGAT 1200
Qy 1254 GAGTACACTTCAGATATTCATGAGTGGGAAAGAGACCAATGTTTGGACAAACATATGC 1313
Db 1201 GAGTACACTTCAGATATTCATGAGTGGGAAAGAGACCAATGTTTGGACAAACATATGC 1260
Qy 1314 GAGATTCCTCTCTAGCTCTCTATTTATCTTGAGCTTGCTCTGAGCTCAGC 1373
Db 1261 GAGGATTCCTCTCTAGCTCTCTATTTATCTTGAGCTTGCTCTGAGCTCAGC 1320
Qy 1374 ACTAGATGAGTTTAAAGCTGAAAATGAGGAAAATTCACACTATTCACCCAGTTGCT 1433
Db 1321 ACTAGATGAGTTTAAAGCTGAAAATGAGGAAAATTCACACTATTCACCCAGTTGCT 1380
Qy 1434 ACCATCTCAGCTACCTCAACCAAGGCTCTCTGTTCCACCGGATCACCACTGATGAT 1493
Db 1381 ACCATCTCAGCTACCTCAACCAAGGCTCTCTGTTCCACCGGATCACCACTGATGAT 1440
Qy 1494 GCATTGTCAAAGCAGCGTCGAATGCTGAAAAATATAGAGGCTTGTTGGATTGGCC 1553
Db 1441 GCATTGTCAAAGCAGCGTCGAATGCTGAAAAATATAGAGGCTTGTTGGATTGGCC 1500
Qy 1554 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1586
Db 1501 CCAGAGATTAACATGATTTCTCGAGTACAAAGTGA 1533
XX
```

RESULT 7

ADQ14494
ID ADQ14494 standard; cDNA; 1533 BP.

AC ADQ14494;

XX 23-SEP-2004 (first entry)

DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #1.

XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;

KM myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;

KW inorganic phosphate; mutant.

XX Glycine max.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 1..1533

FT /tag= a

FT /product= "Mutant soybean myo-inositol 1-phosphate

FT synthase #1"

FT mutation

FT replace(1241,G)

FT /tag= b

XX US2004128713-A1.

XX 01-JUL-2004.

XX 21-NOV-2003; 2003US-00718952.

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX 26-APR-1999; 99US-0029315.

XX 11-MAR-2003; 2002US-00025003.

XX (HITZ/) HITZ W D.

XX (SEBA/) SEBASTIAN S A.

XX (GRAC/) GRACE D J.

XX (STREIT/) STREIT L G.

XX

PI Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX MPI, 2004-533135/51.
DR P-PSDB; ADQ14495.
XX
XX New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX Claim 10; SEQ ID NO 5; 48bp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 87.0%; Score 1531.4; DB 12; Length 1533;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 54 ATGTTTCATGAGAAATTTTAAAGTTGATGATGTCCTTAATGCAATACACCGAGCTGAGATT 113
Db 1 ATGTTTCATGAGAAATTTTAAAGTTGATGATGTCCTTAATGCAATACACCGAGCTGAGATT 60
Qy 114 CAGTCCGTTGACACATGACGAAACCGAATCTGTTACAGAAACGAAATGGACCTAT 173
Db 61 CAGTCCGTTGACACATGACGAAACCGAATCTGTTACAGAAACGAAATGGACCTAT 120
Qy 174 CAGTGGATTTGCAAAACCAATCTGTCAATACGAATTTAAACCAACATTCATGTTCT 233
Db 121 CAGTGGATTTGCAAAACCAATCTGTCAATACGAATTTAAACCAACATTCATGTTCT 180
Qy 234 AAATTAGGGGTAAATGCTGTTGGGTTGGGGTGAAACACAGGCTCAACCGTGGT 293
Db 181 AAATTAGGGGTAAATGCTGTTGGGTTGGGGTGAAACACAGGCTCAACCGTGGT 240
Qy 294 GTTATTGCTAACCGAGAGGGCATTTCAAGGGCTACAAAGGACAAAGTTCAACAGCAAT 353
Db 241 GTTATTGCTAACCGAGAGGGCATTTCAAGGGCTACAAAGGACAAAGTTCAACAGCAAT 300
Qy 354 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCGATGGTGGTCTCTCCAGGAGAGGAA 413
Db 301 TACTTTGGCTCCCTCAACCAAGCTCAGCTATCCGATGGTGGTCTCTCCAGGAGAGGAA 360
Qy 414 ATCTATGCCCATTTCAAGAGCTGCTCTTCAATGTTTAAACCTGACGACATTTGTGGG 473
Db 361 ATCTATGCCCATTTCAAGAGCTGCTCTTCAATGTTTAAACCTGACGACATTTGTGGG 420
Qy 474 GGATGGGATATCAGCAACATGAACCTGAGCTGATGTCATGCGCAGGAGCAAGGATGTTGAC 533
Db 421 GGATGGGATATCAGCAACATGAACCTGAGCTGATGTCATGCGCAGGAGCAAGGATGTTGAC 480
Qy 534 ATCGATTTGACAGAACAGTTGAGGCTTACATGGAATCATGCTTCACTCCCGGAATC 593
Db 481 ATCGATTTGACAGAACAGTTGAGGCTTACATGGAATCATGCTTCACTCCCGGAATC 540
```


QY	594	TATAGACCCGGATTTTCATTGCTGCGCCAAACAAAGAGAGCGGTGCAACAACGTATCAAGAGGC	655
Db	541	TATAGACCCGGATTTTCATTGCTGCGCCAAACAAAGAGAGCGGTGCAACAACGTATCAAGAGGC	600
QY	654	ACAAAGCAAGAGCAAGTTCACAAATCATCAAAGACATCAAGGCGTTTAAAGAACCAAC	713
Db	601	ACAAAGCAAGAGCAAGTTCACAAATCATCAAAGACATCAAGGCGTTTAAAGAACCAAC	660
QY	714	AAAGTGCACAAAGGTGTGTACTGTGGACTGCGCAACACAGAGAGGTACGTAAATTTGGTT	773
Db	661	AAAGTGCACAAAGGTGTGTACTGTGGACTGCGCAACACAGAGAGGTACGTAAATTTGGTT	720
QY	774	GTGGGCGCTTAATGACACCATGAGAAATCTTTGGCTGCTGGAGACAGAAATGAGCGTGAG	833
Db	721	GTGGGCGCTTAATGACACCATGAGAAATCTTTGGCTGCTGGAGACAGAAATGAGCGTGAG	780
QY	834	ATTTCCTCCTTCACACCTGTGTATGCCATTGCTGTATGAGAAATGTTCCTTTCAATTAAT	893
Db	781	ATTTCCTCCTTCACACCTGTGTATGCCATTGCTGTATGAGAAATGTTCCTTTCAATTAAT	840
QY	894	GGAAGCCCTCAGAACACTTTTGTGTACAGGGCTGATTTGATCTTTGCCATGTGGAGAACACT	953
Db	841	GGAAGCCCTCAGAACACTTTTGTGTACAGGGCTGATTTGATCTTTGCCATGTGGAGAACACT	900
QY	954	TTGATTTGTGAGAGATGACTTCAGAGGTGGTCAGACCAAAATGAATCTGTGTGGTTGAT	1013
Db	901	TTGATTTGTGAGAGATGACTTCAGAGGTGGTCAGACCAAAATGAATCTGTGTGGTTGAT	960
QY	1014	TTTCCTTGTGGGGCGTGATTAAGGCCAACATCTATATGTCACTTAACAACCAATCTGGGAAAC	1073
Db	961	TTTCCTTGTGGGGCGTGATTAAGGCCAACATCTATATGTCACTTAACAACCAATCTGGGAAAC	1020
QY	1074	AATATATGTATGAATCTTTGGGCTCCAAACTTTCCGTTCCAGAGAAATCTCCAAAGAC	1133
Db	1021	AATATATGTATGAATCTTTGGGCTCCAAACTTTCCGTTCCAGAGAAATCTCCAAAGAC	1080
QY	1134	AACGTTGTGATGATATGATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGGAAATCTCA	1193
Db	1081	AACGTTGTGATGATATGATGCTCAACAGCAATGCCATCTCTATAGAGCTGTGGAAATCTCA	1140
QY	1194	GACCATGTTGTTGTTATTAAGTATGTGCTTAACGTAGGGGACAGCAGAGAGCCATGGAT	1253
Db	1141	GACCATGTTGTTGTTATTAAGTATGTGCTTAACGTAGGGGACAGCAGAGAGCCATGGAT	1200
QY	1254	GAGTACACTTCAGAGATATATCATGGGTGAAAGAGCAACTGTGTTTTCACAAACAATGC	1313
Db	1201	GAGTACACTTCAGAGATATATCATGGGTGAAAGAGCAACTGTGTTTTCACAAACAATGC	1260
QY	1314	GAGCATTCCTCTTAGCTGCTCTTATTAATCTTGGACTTGGTCTCTTGTAGCTCAGC	1373
Db	1261	GAGCATTCCTCTTAGCTGCTCTTATTAATCTTGGACTTGGTCTCTTGTAGCTCAGC	1320
QY	1374	ACTAGAAATCGAGTTTAAAGCTGAAAAATGAGGAAATTCCTACTACTTCCACCCAGTTGCT	1433
Db	1321	ACTAGAAATCGAGTTTAAAGCTGAAAAATGAGGAAATTCCTACTACTTCCACCCAGTTGCT	1380
QY	1434	ACCATCTCAGGTACTCTCACCAGAGGCTCCTCGGTTCCACGGGGATCAACAGTGGTGAAT	1493
Db	1381	ACCATCTCAGGTACTCTCACCAGAGGCTCCTCGGTTCCACGGGGATCAACAGTGGTGAAT	1440
QY	1494	GCAATTTGCAAAAGCAGCGTGCAATGCTGAGAAAACATTAATGAGGCGCTGTGTGGATTGGCC	1553
Db	1441	GCAATTTGCAAAAGCAGCGTGCAATGCTGAGAAAACATTAATGAGGCGCTGTGTGGATTGGCC	1500
QY	1554	CCAGAGAAATTAATGATATCTTGCAGATTAACAATGCA	1586
Db	1501	CCAGAGAAATTAATGATATCTTGCAGATTAACAATGCA	1533

RESULT 8	
ADS81997	
ID	ADS81997 standard; cDNA; 1533 BP.
XX	

AC	ADS81997,
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	Soybean myo-inositol 1-phosphate synthase cDNA mutant #1.
XX	
KM	Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX	raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
XX	
OS	Glycine max; line LR33.
OS	Synthetic.
XX	
FH	Key
FT	Location/Qualifiers
FT	1..1533
FT	/tag= a
FT	/product= "myo-inositol 1-phosphate synthase"
FT	/replace(1188,G)
FT	/tag= b
XX	
PN	US2003074685-A1.
XX	
PD	17-APR-2003.
XX	
PE	11-MAR-2002; 2002US-00025003.
XX	
XX	08-APR-1997; 97US-00835751.
PR	07-APR-1998; 98WO-US006822.
XX	
XX	
PA	(HITZ/) HITZ W D.
PA	(SEBA/) SEBASTIAN S A.
XX	
PI	Hitz WD, Sebastian SA;
XX	
DR	WPI; 2004-639957/62.
XX	
DR	P-PSDB; ADS81998.
XX	
PT	Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
PT	phosphate synthase, useful for altering raffinose saccharide, sucrose,
PT	phytic acid and inorganic phosphate content of soybean seeds.
XX	
PS	Claim 8; SEQ ID NO 5; 34pp; English.
XX	
XX	The invention relates to an isolated nucleic acid fragment encoding a
CC	soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
CC	phosphate synthase having decreasing capacity for the synthesis for myo-
CC	inositol-1-phosphate. Also included are a chimeric gene (comprising the
CC	nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
CC	complement, subfragment or the complement of the subfragment, operably
CC	linked to, suitable regulatory sequences, where expression of the chimeric
CC	gene results in a decrease in expression of an endogenous or native gene
CC	encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
CC	comprising the chimeric gene (with a heritable phenotype of a seed phytic
CC	acid content of less than 17 micromol/g, a seed content of raffinose plus
CC	stachyose of less than 14.5 micromol/g, and a seed sucrose content of
CC	greater than 200 micromol/g, provided that the plant is not LR33), seeds
CC	from the plant, making a soybean plant with the heritable phenotype
CC	(comprising crossing LR33 or the plant comprising the chimeric gene with
CC	an elite soybean plant and selecting a progeny plant of the cross of
CC	crossing step that has a heritable phenotype as mentioned above), seeds
CC	of soybean plant made by the above method, a soy protein product derived
CC	from seeds of a soybean plant (homozygous for one or more gene encoding a
CC	mutant myo-inositol 1-phosphate synthase having decreased capacity for
CC	the synthesis of myo-inositol 1-phosphate, where the gene confers a
CC	heritable phenotype as mentioned above), and making or producing a
CC	soybean protein product derived from seeds of a soybean plant with a
CC	heritable phenotype as mentioned above. The nucleic acid is useful for
CC	altering raffinose saccharide, sucrose, phytic acid and inorganic
CC	phosphate content of soybean seeds thus leading to valuable and useful
CC	soybean products, since the presence of high concentration of raffinose
CC	oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC	when consumed by humans. The present sequence encodes a mutant myo-
XX	inositol 1-phosphate synthase.
XX	

Sequence 1533 BP; 429 A; 344 C; 367 G; 393 T; 0 U; 0 Other;

Query Match 87.0%; Score 1531.4; DB 13; Length 1533;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 54 ATGTCATCGAGAAATTTTAAGTTGAGTGTCTAATGTGAAGTACCCGAGCTGAGATT 113
    |||
Db 1 ATGTCATCGAGAAATTTTAAGTTGAGTGTCTAATGTGAAGTACCCGAGCTGAGATT 60
    |||

QY 114 CAGTCCGTGTACAACTACGAAACACCGAACTTGTTACGAGAACAGAAATGACCTAT 173
    |||
Db 61 CAGTCCGTGTACAACTACGAAACACCGAACTTGTTACGAGAACAGAAATGACCTAT 120
    |||

QY 174 CAGTGAATGTCAAAACCCAAATCTGTCAATACGAATTTAAACCAACATCCATTTCT 233
    |||
Db 121 CAGTGAATGTCAAAACCCAAATCTGTCAATACGAATTTAAACCAACATCCATTTCT 180
    |||

QY 234 AAATTAGGGGTAAATGCTTGTGGGTGGGTGAAACAAAGGCTCAACCCCTCACGGTGT 293
    |||
Db 181 AAATTAGGGGTAAATGCTTGTGGGTGGGTGAAACAAAGGCTCAACCCCTCACGGTGT 240
    |||

QY 294 GTTATTGCTAACCGAGAGGGCATTTTCATGCGCTACAAAGGACAAAGATTCAACAGCAAT 353
    |||
Db 241 GTTATTGCTAACCGAGAGGGCATTTTCATGCGCTACAAAGGACAAAGATTCAACAGCAAT 300
    |||

QY 354 TACTTTGGCTCCTTCAACCCAGCTTCACTATCCGAGTTGGTCTTCCAGGGAGAGAA 413
    |||
Db 301 TACTTTGGCTCCTTCAACCCAGCTTCACTATCCGAGTTGGTCTTCCAGGGAGAGAA 360
    |||

QY 414 ATCTATGCCCCATTCAGAGGCTGCTTCCAAATGTTTAACTTACCTACGACATGTGTTGG 473
    |||
Db 361 ATCTATGCCCCATTCAGAGGCTGCTTCCAAATGTTTAACTTACCTACGACATGTGTTGG 420
    |||

QY 474 GGAATGGATATACGAACAATGAACCTGCTGATGCCATGCGCAAGGCAAGTGTGAC 533
    |||
Db 421 GGAATGGATATACGAACAATGAACCTGCTGATGCCATGCGCAAGGCAAGTGTGAC 480
    |||

QY 534 ATCGATTGTGACAGAGCATGTTAGGCTTTCATGGAATCCATGCTTCCACTCCCCGGAATC 593
    |||
Db 481 ATCGATTGTGACAGAGCATGTTAGGCTTTCATGGAATCCATGCTTCCACTCCCCGGAATC 540
    |||

QY 594 TATGACCCGGATTTCAATGCTGCCAACAGAGAGGTGSCAAACACGTCATCAAGGGC 653
    |||
Db 541 TATGACCCGGATTTCAATGCTGCCAACAGAGAGGTGSCAAACACGTCATCAAGGGC 600
    |||

QY 654 ACAAGCAAGAGAGCATTTCAAAATCATCAAGACATCAAGGCGTTTAAAGAACACC 713
    |||
Db 601 ACAAGCAAGAGAGCATTTCAAAATCATCAAGACATCAAGGCGTTTAAAGAACACC 660
    |||

QY 714 AAAGTGACAGAGTGTGTTGATCTGTGACTGCCAACACAGAGAGGTACAGTAATTTGGTT 773
    |||
Db 661 AAAGTGACAGAGTGTGTTGATCTGTGACTGCCAACACAGAGAGGTACAGTAATTTGGTT 720
    |||

QY 774 GTGGGCTTAAATGACACCATGGAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 833
    |||
Db 721 GTGGGCTTAAATGACACCATGGAATCTTGGCTGCTGTGACAGAAATGAGGCTGAG 780
    |||

QY 834 ATTTCTCTTCCACTTGTATGCGCATTTGTTGATGGAATAATGTTCTTTCATTAAAT 893
    |||
Db 781 ATTTCTCTTCCACTTGTATGCGCATTTGTTGATGGAATAATGTTCTTTCATTAAAT 840
    |||

QY 894 GGAAGCCTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTTCATTCGAGGAACACT 953
    |||
Db 841 GGAAGCCTCAGAACCTTTTGTACAGAGGCTGATTTGATCTTTCATTCGAGGAACACT 900
    |||

QY 954 TTGATTTGGTGGAGATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 1013
    |||
Db 901 TTGATTTGGTGGAGATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGAT 960
    |||

QY 1014 TTCTTTGTGGGGCTGTATCAAGCCAACTATATGTAGTACCAACCATCTGGGAAC 1073
    |||
Db 961 TTCTTTGTGGGGCTGTATCAAGCCAACTATATGTAGTACCAACCATCTGGGAAC 1020
    |||

```

```

QY 1074 AATGATGTATGAATCTTTCGGCTCCACAAACTTTCGGTTCCAAAGGAATCTCCAGAGC 1133
    |||
Db 1021 AATGATGTATGAATCTTTCGGCTCCACAAACTTTCGGTTCCAAAGGAATCTCCAGAGC 1080
    |||

QY 1134 AACGTTGTGATGATATGTGTCAACAGCAATGCGCATCTCTATGAGCGCTGTGAATCCA 1193
    |||
Db 1081 AACGTTGTGATGATATGTGTCAACAGCAATGCGCATCTCTATGAGCGCTGTGAATCCA 1140
    |||

QY 1194 GACCATGTTGTTTATTAAGTATGTGCTTACGTAAGGAGACAGCAAGAGCCATGAT 1253
    |||
Db 1141 GACCATGTTGTTTATTAAGTATGTGCTTACGTAAGGAGACAGCAATGAGCCATGAT 1200
    |||

QY 1254 GAGTACCTTCAGAGATATTCATAGGTTGGAAGAGACCAATGTTTTCACACATGC 1313
    |||
Db 1201 GAGTACCTTCAGAGATATTCATAGGTTGGAAGAGACCAATGTTTTCACACATGC 1260
    |||

QY 1314 GAGGATTCCTCTTAGCTGCTCCTATTTATCTTGAATGATGATGATGATGATGATGAT 1373
    |||
Db 1261 GAGGATTCCTCTTAGCTGCTCCTATTTATCTTGAATGATGATGATGATGATGATGAT 1320
    |||

QY 1374 ACTAGATCGAGTTTAAAGCTGAAATGAGGAAATTCACCTCATTTCCACCAAGTTGCT 1433
    |||
Db 1321 ACTAGATCGAGTTTAAAGCTGAAATGAGGAAATTCACCTCATTTCCACCAAGTTGCT 1380
    |||

QY 1434 ACCATCTCAGCTACTCTCAACAGGCTCTCTGTGTTCCACCGGATACACAGTGTGAAT 1493
    |||
Db 1381 ACCATCTCAGCTACTCTCAACAGGCTCTCTGTGTTCCACCGGATACACAGTGTGAAT 1440
    |||

QY 1494 GCATTGTCAAGAGGCTGCAATGCTGAAATCATTAATGAGGCTTGTGTGATTTGGCC 1553
    |||
Db 1441 GCATTGTCAAGAGGCTGCAATGCTGAAATCATTAATGAGGCTTGTGTGATTTGGCC 1500
    |||

QY 1554 CCAGAGATTAACATGATTTCTCGATACAAAGTGA 1586
    |||
Db 1501 CCAGAGATTAACATGATTTCTCGATACAAAGTGA 1533
    |||

```

RESULT 9

ADQ14504
ID ADQ14504 standard; cDNA, 1533 BP.

XX ADQ14504;

XX 23-SBP-2004 (first entry)

XX Wild type soybean myo-inositol 1-phosphate synthase cDNA #2.

XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;

KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;

XX inorganic phosphate.

XX Glycine max.

XX

XX

XX US2004128713-A1.

XX 01-JUL-2004.

XX 21-NOV-2003; 2003US-00718952.

XX

XX 08-APR-1997; 97US-00835751.

XX 07-APR-1998; 98WO-US006822.

XX 26-APR-1999; 99US-00299315.

XX 11-MAR-2002; 2002US-00025003.

XX (HITZ/) HITZ W D.

PA (SEBA/) SEBASTIAN S A.

PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
XX
XX Hitz WD, Sebastian SA, Grace DJ, Streit LG;
XX WPI: 2004-53335/51.
XX P-PSDB; ADQ14505.
XX

PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX

PS Claim 4; SEQ ID NO 15; 48bp; English.

CC The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a wild type
CC soybean myo-inositol 1-phosphate synthase polypeptide of the invention.
XX

XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 83.3%; Score 1465.8; DB 12; Length 1533;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1491; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 54 ATGTCATGAGAAATTTTAAAGTTGAGTCTTAATGTGAATGACCGAGACTGAGATT 113
DB 1 ATGTCATGAGAAATTTTAAAGTTGAGTCTTAATGTGAATGACCGAGACTGAGATT 60
QY 114 CAGTCGCTGATACAACTACGAAACCAAGCACTGTTACGAGAAACGAATGSCACTAT 173
DB 61 CAGTCGCTGATACAACTACGAAACCAAGCACTGTTACGAGAAACGAATGSCACTAT 120
QY 174 CAGTGATTTGTCAAAACCAATCTGTCAAAATGCAATTTAAACCAACATTCATGTTCT 233
DB 121 CAGTGATTTGTCAAAACCAATCTGTCAAAATGCAATTTAAACCAACATTCATGTTCA 180
QY 234 AAATTTAGGGGTAATGTTTGGGTTGGGGTGGGAAACAAAGGCTCAACCTTCACCGTGT 293
DB 181 AAATTTAGGGGTAATGTTTGGGTTGGGGTGGGAAACAAAGGCTTCACCTTCACCGTGT 240
QY 294 GTTATTTGCTAACGAGAGGCACTTTTCATGGGCTACAAAGCAAGATTCGAACAGCAAT 353
DB 241 GTTATTTGCTAACGAGAGGCACTTTTCATGGGCTACAAAGCAAGATTCGAACAGCAAT 300
QY 354 TACTTTGGCTCCCTCAACCAAGGCTCAGTATTCGAGTTGGGTCCTTCAGGAGAGAGAA 413
DB 301 TACTTTGGCTCCCTCAACCAAGGCTCAGTATTCGAGTTGGGTCCTTCAGGAGAGAGAA 360
QY 414 ATCTATGCCCATTTCAAGGCTGCTTCCAAATGTTTAACTTCAGACATTTGTTGGG 473
DB 361 ATCTATGCCCATTTCAAGGCTGCTTCCAAATGTTTAACTTCAGACATTTGTTGGG 420
QY 474 GGATGGATATATGAGCAATGAACCTGGCTGATGACATGGGCAAGGCAAGGTTTTCAC 533
DB 421 GGATGGATATATGAGCAATGAACCTGGCTGATGACATGGGCAAGGCAAGGTTTTCAC 480
QY 534 ATCGATTTCAGAAAGCAATTGAGGCTTACATGAATTCATGCTTCCCTCCCGGAATC 593

DB 481 ATCGATTTCAGAAAGCAATTGAGGCTTACATGAATTCATGTTCCACTCCCGGAATC 540
QY 594 TATGACCCGGGATTTTATGTTGCGCCCAACGAGGAGCGTGCACAAAGTCTCAAGGAC 653
DB 541 TACGACCCGGGATTTTATGTTGCGCCCAACGAGGAGCGTGCACAAAGTCTTAAAGGAC 600
QY 654 ACAAGCAAGAGCAAGTTCAACAAATCATCAAGACATCAAGGCTTTAAGAAACCAAC 713
DB 601 ACAAGCAAGAGCAAGTTCAAGAAATCATCAAGACATCAAGGCTTTAAGAAACCAAC 660
QY 714 AAATGGACAAAGGTGTTGATCTGGGACTGCCCAACAGAGAGGTACAGTAATTTGGTT 773
DB 661 AAATGGACAAAGGTGTTGCTGTTGACTGCCCAACAGAGAGGTATAGCAATTTGGTT 720
QY 774 GTGGGCTTTATGACACCATGAGAGAAATCTTGGCTGCTGTGAGACAGAAATGAGGCTGAG 833
DB 721 GTAGGCTTTATGACACCATGAGAGAAATCTTGGCTGCTGTGAGACAGAAATGAGGCTGAG 780
QY 834 ATTTCTCTTCACCTTGTATGSCATTTGCTGTGTTATGAGAAATGTTCTTTCAATTAAT 893
DB 781 ATTTCTCTTCACCTTGTATGSCATTTGCTGTGTTATGAGAAATGTTCTTTCAATTAAT 840
QY 894 GGAAGCCTTCAGAACACTTTTGTATCCAGGCGTATGATCTTTGCCATGCGAGAACT 953
DB 841 GGAAGCCTTCAGAACACTTTTGTATCCAGGCGTATGATCTTTGCCATGCGAGAACT 900
QY 954 TTGATTTGGAGATGACCTTCAAGAGTGTGACAGACCAAAATGAATCTGTGTTGGTTGAT 1013
DB 901 TTGATTTGGAGATGACCTTCAAGAGTGTGACAGACCAAAATGAATCTGTGTTGGTTGAT 960
QY 1014 TTCTCTTGGGGGCTGTATGACGCAACATCTATAGTCACTTACCAACCATCTGGGAAAC 1073
DB 961 TTCTCTTGGGGGCTGTATGACGCAACATCTATAGTCACTTACCAACCATCTGGGAAAC 1020
QY 1074 AATGATGATGAATCTTTTGGCTTCACAACTTTCCGTTCCAGGAATCTCCAAAGAC 1133
DB 1021 AATGATGATGAATCTTTGGCTTCACAACTTTCCGTTCCAGGAATCTCCAAAGAC 1080
QY 1134 AACGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
DB 1081 AACGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1194 GACCATGTTGTTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1253
DB 1141 GACCATGTTGTTGTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1254 GAGTACACTTCAGAGATATTCATGAGGTGAAAGAGACATCTTTTGTGCAACAACATGTC 1313
DB 1201 GAGTACACTTCAGAGATATTCATGAGGTGAAAGAGACATCTTTTGTGCAACAACATGTC 1260
QY 1314 GAGGATTCCTCTTGAAGTCTCTATTAATCTTGAAGTCTCTTGAAGTCTCTTGAAGTCTCAAG 1373
DB 1261 GAGGATTCCTCTTGAAGTCTCTATTAATCTTGAAGTCTCTTGAAGTCTCTTGAAGTCTCAAG 1320
QY 1374 ACTAGAAATGAGTTTAAAGCTGAAATGAGGAAATTCATCTCACTATTCACCCAGTGTCT 1433
DB 1321 ACTAGAAATGAGTTTAAAGCTGAAATGAGGAAATTCATCTCACTATTCACCCAGTGTCT 1380
QY 1434 ACCATCTCAGTATCTTCAACCAAGGCTCTCTGTTTCCACCGGGTACACAGTGTGAT 1493
DB 1381 ACCATCTCAGTATCTTCAACCAAGGCTCTCTGTTTCCACCGGGTACACAGTGTGAT 1440
QY 1494 GGATTTGCAAAAGCAGCGTGAATGCTGGAAGAAACATTAATGAGGAGCTTGTGATTTGGCC 1553
DB 1441 GGATTTGCAAAAGCAGCGTGAATGCTGGAAGAAACATTAATGAGGAGCTTGTGATTTGGCC 1500
QY 1554 CCAGAGAAATTAATCATGATTTCTGAGTACAAAGTGA 1586
DB 1501 CCAGAGAAATTAATCATGATTTCTGAGTACAAAGTGA 1533

RESULT 10

ADQ14502
ID ADQ14502 standard; cDNA; 1533 BP.
XX
AC ADQ14502;
XX
DT 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #4.
XX
KM Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
XX inorganic phosphate; mutant.
XX
OS Glycine max.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "Mutant soybean myo-inositol 1-phosphate
FT synthase #4"
XX
PN US2004128713-A1.
XX
PD 01-JUL-2004.
XX
PP 21-NOV-2003; 2003US-00718952.
XX
PR 08-APR-1997; 97US-00835751.
PR 07-APR-1998; 98WO-US006822.
PR 26-APR-1999; 99US-0029315-
PR 11-MAR-2002; 2002US-00025003.
XX
PA (HITZ/) HITZ W D.
PA (SEBA/) SEBASTIAN S A.
PA (GRAC/) GRACE D J.
PA (STRE/) STREIT L G.
XX
PI Hitz WD, Sebastian SA, Grace DJ, Streit LG,
XX
DR WPI; 2004-533135/51.
DR P-PSDB; ADQ14503.
XX
XX
PT New nucleic acid fragment encoding myo-inositol 1-phosphate synthase,
PT useful for producing plants with decreased raffinose, stachyose, and
PT phytic acid and increased sucrose, leading to valuable and useful soybean
PT products.
XX
XX
PS Example 8; SEQ ID NO 13; 48pp; English.
XX
XX The invention relates to a nucleic acid fragment encoding a soybean myo-
CC inositol 1-phosphate synthase or a mutant myo-inositol 1-phosphate
CC synthase having a decreased capacity for the synthesis of myo-inositol 1-
CC phosphate. The invention also relates to a chimeric gene operably linked
CC to suitable regulatory sequences, where expression of the chimeric gene
CC results in a decrease in expression of an endogenous or native gene
CC encoding a soybean myo-inositol 1-phosphate synthase, a soybean plant
CC comprising the chimeric gene, a method of making the soybean plant, a
CC seed of the soybean plant, a soy protein product derived from the plant,
CC processing of soybean seeds, a method of making or producing a soy
CC protein product and a method of using a soybean plant homozygous for at
CC least one gene encoding a mutant myo-inositol 1-phosphate synthase having
CC decreased capacity for the synthesis of myo-inositol 1-phosphate. The
CC nucleic acid fragment and methods are useful for producing plants with
CC decreased raffinose, stachyose and phytic acid content and increased
CC sucrose and inorganic phosphate content, leading to valuable and useful
CC soybean products. This sequence represents cDNA encoding a mutant soybean
CC myo-inositol 1-phosphate synthase polypeptide of the invention.
XX
SQ Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;

Query Match 83.3%; Score 1465.8; DB 12; Length 1533;
Best Local Similarity 97.3%; Pred. No. 0;

	Matches 1491;	Conservative 0;	Mismatches 42;	Indels 0;	Gaps 0;
QY	54	ATGTTGATGAGAAATTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGATT			113
DB	1	ATGTTGATGAGAAATTTAAAGTTGAGTGTCTTAATGTGAAGTACACCGAGCTGAGATT			60
QY	114	CAGTCGGTGTACAACTACGAAACCAACCGAACTTTGTCAGAGAACGAAATGGCACTTAT			173
DB	61	CAGTCGGTGTACAACTACGAAACCAACCGAACTTTGTCAGAGAACGAAATGGCACTTAT			120
QY	174	CAGTGATTTGTCAAAACCAATCTGTCAATATGCAATTTAAACCAACATCCATGTTTCT			233
DB	121	CAGTGATTTGTCAAAACCAATCTGTCAATATTTAAACCAACATCCATGTTTCTCA			180
QY	234	AAATTAAGGGTAAATGTTGTGGTTGGGTGGAACAACGCTCAACCTCACCGTGGT			293
DB	181	AAATTAAGGGGTAAATGTTGTGGTTGGGTGGAACAACGCTCTACCTCACCGTGGT			240
QY	294	GTTATTTGCTAACCGAGGGCAATTTTCATGGGCTACAAAGACAAAGATTCAACAGCCAA			353
DB	241	GTTATTTGCTAACCGAGGGCAATTTTCATGGGCTACAAAGACAAAGATTCAACAGCCAA			300
QY	354	TACTTGGCTCCCTCACCAAGCTCAGCTATCCGAGTTGGTCTTCCAGGGAGAGAA			413
DB	301	TACTTGGCTCCCTCACCAAGCTCAGCTATTCAGATTGATTCCTTCCAGGGAGAGAA			360
QY	414	ATCTATGCCCCATTCAAGCCTGCTTCCATATGTTTAACTCTGACGACATTTGTGTTGG			473
DB	361	ATCTATGCCCCATTCAAGCCTGCTTCCATATGTTTAACTCTGACGACATTTGTGTTGG			420
QY	474	GGATGGGATATACGACAACTGAACCTGTGATGTCATGCGCCAGGGCAAAAGGTTTGAC			533
DB	421	GGATGGGATATACGACAACTGAACCTGTGATGTCATGCGCCAGGGCAAAAGGTTTGAC			480
QY	534	ATCGATTTGCAAGAGAGTTGAGGCTTACATGAAATCATGCTTCACTCCCGGAATC			593
DB	481	ATCGATTTGCAAGAGAGTTGAGGCTTACATGAAATCATGCTTCACTCCCGGAATC			540
QY	594	TATGACCCGGAATTTCAATGCTGCTCCAAACCAAGAGAGCTGCCAACACGTCATCAAGGC			653
DB	541	TATGACCCGGAATTTCAATGCTGCTCCAAACCAAGAGAGCTGCCAACACGTCATCAAGGC			600
QY	654	ACAAAGCAAGACCAAGTTTCAAAATATCAAAAGATCAAAAGCCGTTTAAAGAACGACAC			713
DB	601	ACAAAGCAAGACCAAGTTTCAAAATATCAAAAGATCAAAAGCCGTTTAAAGAACGACAC			660
QY	714	AAAGTGACCAAGGTGTTGTACTGTGACGTGCAACACAGAGAGGATAGTAAATTTGGTT			773
DB	661	AAAGTGACCAAGGTGTTGTACTGTGACGTGCAACACAGAGAGGATAGTAAATTTGGTT			720
QY	774	GTTGGCTTTAATGACACCATGAGAAATCTTTGGCTGTGTGAGACGAATAGAGCTGAG			833
DB	721	GTTGGCTTTAATGACACCATGAGAAATCTTTGGCTGTGTGAGACGAATAGAGCTGAG			780
QY	834	ATTTCTCTTCCACTTGTATGCCATGTGTGTATATGAAAATGTTCTTTCAATAT			893
DB	781	ATTTCTCTTCCACTTGTATGCCATGTGTGTATATGAAAATGTTCTTTCAATAT			840
QY	894	GGAGGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAACACT			953
DB	841	GGAGGCCCTCAGAACACTTTTGTACAGAGGCTGATTTGATCTTGCCATGCGAGAACACT			900
QY	954	TTGATTTGTGAGATGACTTCAAGAGTGTACAGACCAAAATGTAATCTGTGTTGTTGAT			1013
DB	901	TTGATTTGTGAGATGACTTCAAGAGTGTACAGACCAAAATGTAATCTGTGTTGTTGAT			960
QY	1014	TTCTTTGGGGGCTGTATACAGCAACCTTAATAGTCAAGTTAACCAATTTGGGAAAC			1073
DB	961	TTCTTTGGGGGCTGTATACAGCAACCTTAATAGTCAAGTTAACCAATTTGGGAAAC			1020
QY	1074	AATGATGATGATGAATCTTGGGCTCCAAACCTTCCGTTCCAGAGAAATCTCAAGAGC			1133
DB	1021	AATGATGATGATGAATCTTGGGCTCCAAACCTTCCGTTCCAGAGAAATCTCAAGAGC			1080

QY 534 ATGATTTGAGAGCAGTTGAGGCTTACATGAAATCCATGCTCCACTCCCCGGATC 593
 DB 481 ATGATTTGAGAGCAGTTGAGGCTTACATGAAATCCATGCTCCACTCCCCGGATC 540
 QY 594 TATGACCCGGATTTCAATGCTGCAACCAAGAGAGGCTCCAAACATCATCAAGGCG 653
 DB 541 TAGGACCCGGATTTCAATGCTGCAACCAAGAGAGGCTCCAAACATCATCAAGGCG 600
 QY 654 ACAAGCAGAGCAAGTTCAACAATCATCAAGAGCATCAAGGCTTTAAGAGCCACC 713
 DB 601 ACAAGCAGAGCAAGTTCAACAATCATCAAGAGCATCAAGGCTTTAAGAGCCACC 660
 QY 714 AAAGTGACAGAGGTTGTTACTGCTGCAACCAAGAGGTTAAGTAATTTGGTT 773
 DB 661 AAAGTGACAGAGGTTGTTACTGCTGCAACCAAGAGGTTAAGTAATTTGGTT 720
 QY 774 GTGGGCTTAAATGACACATGAGAAATCTTGGCTGCTGAGACAGAAATGAGCTGAG 833
 DB 721 GTAGGCTTAAATGACACATGAGAAATCTTGGCTGCTGAGACAGAAATGAGCTGAG 780
 QY 834 ATTTCTCTTCCACCTTGTATGCAATGCTGTTATGGAATAATGTTCTTCTTCAATTA 893
 DB 781 ATTTCTCTTCCACCTTGTATGCAATGCTGTTATGGAATAATGTTCTTCTTCAATTA 840
 QY 894 GGAAGCCTCAGAACCTTTGTACAGAGGCTGATTTGATCTTGCATCCGAGGAACACT 953
 DB 841 GGAAGCCTCAGAACCTTTGTACAGAGGCTGATTTGATCTTGCATCCGAGGAACACT 900
 QY 954 TTGATTTGAGAGATGATCTTCAAGAGTGTCAAGCAAAATGAATCTGTGTTGAT 1013
 DB 901 TTGATTTGAGAGATGATCTTCAAGAGTGTCAAGCAAAATGAATCTGTGTTGAT 960
 QY 1014 TTCTTTGAGGAGCTGTATCAAGCCAACTTATATGTCATTTACCAACCTTGGAAAC 1073
 DB 961 TTCTTTGAGGAGCTGTATCAAGCCAACTTATATGTCATTTACCAACCTTGGAAAC 1020
 QY 1074 AATGATGATATGATCTTTCGAGCTCCAAACCTTTCGTTCCAAAGAAATCTCCAGAGC 1133
 DB 1021 AATGATGATATGATCTTTCGAGCTCCAAACCTTTCGTTCCAAAGAAATCTCCAGAGC 1080
 QY 1134 AACGTTGATGATATGTCACACAGCAATGCCATCTTATAGAGCTGTGAACATCCA 1193
 DB 1081 AACGTTGATGATATGTCACACAGCAATGCCATCTTATAGAGCTGTGAACATCCA 1140
 QY 1194 GACCATGTTGTTTATTAAGTATGTCCTTACGTAAGGAGCAGCAAGAGCCATGAT 1253
 DB 1141 GACCATGTTGTTTATTAAGTATGTCCTTACGTAAGGAGCAGCAAGAGCCATGAT 1200
 QY 1254 GAGTACCTTCAAGATATTCATGAGGTGGAAGAGACCATTTGTTGACACACATGC 1313
 DB 1201 GAGTACCTTCAAGATATTCATGAGGTGGAAGAGACCATTTGTTGACACACATGC 1260
 QY 1314 GAGGATTCCTCTTACCTGCTCTATTAATCTTGAATGCTGCTTCTGCTGAGCTGAGC 1373
 DB 1261 GAGGATTCCTCTTACCTGCTCTATTAATCTTGAATGCTGCTTCTGCTGAGCTGAGC 1320
 QY 1374 ACTAGATGAGTTTAAAGCTGAAATGAGGAAAAATTCACATCTATTCACACCATGTTGCT 1433
 DB 1321 ACTAGATGAGTTTAAAGCTGAAATGAGGAAAAATTCACATCTATTCACACCATGTTGCT 1380
 QY 1434 ACCATCTCAGCTACCTCAACCAAGGCTCTCTGCTTCCACCGGCTACACAGTGGTGAAT 1493
 DB 1381 ACCATCTCAGCTACCTCAACCAAGGCTCTCTGCTTCCACCGGCTACACAGTGGTGAAT 1440
 QY 1494 GATTTGTCAAAGCAGCTCAATGCTGAAACCATATGAGGCTTGTGTGATTTGGCC 1553
 DB 1441 GATTTGTCAAAGCAGCTCAATGCTGAAACCATATGAGGCTTGTGTGATTTGGCC 1500
 QY 1554 CCAGAGATTAACATGATTTCTCGATCAACAGTGA 1586
 DB 1501 CCAGAGATTAACATGATTTCTCGATCAACAGTGA 1533

RESULT 12
 ADS82003
 ID ADS82003 standard; cDNA, 1533 BP.
 XX
 AC ADS82003;
 XX
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Soybean myo-inositol 1-phosphate synthase cDNA, wild-type 4.
 XX
 KM Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
 KM raffinose; stachyose; sucrose; inorganic phosphate; flatulence.
 XX
 OS Glycine max; line 29018BP03.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1533
 FT /*tag= a
 FT /product= "myo-inositol 1-phosphate synthase"
 XX
 PN US2003074665-A1.
 PD
 XX 17-APR-2003.
 PF 11-MAR-2002; 2002US-00025003.
 XX
 PR 08-APR-1997; 97US-00835751.
 PR 07-APR-1998; 98WO-US006822.
 XX
 PA (HITZ/) HITZ W D.
 PA (SEBA/) SEBASTIAN S A.
 XX
 PI Hitz WD, Sebastian SA;
 XX
 DR WPI; 2004-63957/62.
 DR P-PSDB; ADS82004.
 XX
 PT Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-
 PT phosphate synthase, useful for altering raffinose saccharide, sucrose,
 PT phytic acid and inorganic phosphate content of soybean seeds.
 XX
 PS Example 8; SEQ ID NO 13; 34pp; English.
 XX
 CC The invention relates to an isolated nucleic acid fragment encoding a
 CC soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-
 CC phosphate synthase having decreasing capacity for the synthesis for myo-
 CC inositol-1-phosphate. Also included are a chimeric gene (comprising the
 CC nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its
 CC complement, subfragment or the complement of the subfragment, operably
 CC linked to suitable regulatory sequences, where expression of the chimeric
 CC gene results in a decrease in expression of an endogenous or native gene
 CC encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant
 CC comprising the chimeric gene (with a heritable phenotype of a seed phytic
 CC acid content of less than 17 micromol/g, a seed content of raffinose plus
 CC stachyose of less than 14.5 micromol/g, and a seed sucrose content of
 CC greater than 200 micromol/g, provided that the plant is not LR33), seeds
 CC from the plant, making a soybean plant with the heritable phenotype
 CC (comprising crossing LR33 or the plant comprising the chimeric gene with
 CC an elite soybean plant and selecting a progeny plant of the cross of
 CC crossing step that has a heritable phenotype as mentioned above), seeds
 CC of soybean plant made by the above method, a soy protein product derived
 CC from seeds of a soybean plant (homozygous for one or more gene encoding a
 CC mutant myo-inositol 1-phosphate synthase having decreased capacity for
 CC the synthesis of myo-inositol 1-phosphate, where the gene confers a
 CC heritable phenotype as mentioned above), and making or producing a
 CC soybean protein product derived from seeds of a soybean plant with
 CC heritable phenotype as mentioned above. The nucleic acid is useful for
 CC altering raffinose saccharide, sucrose, phytic acid and inorganic
 CC phosphate content of soybean seeds thus leading to valuable and useful
 CC soybean products, since the presence of high concentration of raffinose
 CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
 CC when consumed by humans. The present sequence encodes a wild-type myo-
 CC inositol 1-phosphate synthase.

```

XX Sequence 1533 BP; 429 A; 341 C; 371 G; 392 T; 0 U; 0 Other;
SQ Query Match 83.3%; Score 1465.8; DB 13; Length 1533;
Beet Local Similarity 97.3%; Pred. No. 0;
Matches 1491; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

OY 54 ATGTCATCGAGAAATTTTAAGGTGAGTGTCTTAATGTGAATGACCGAGACTGAGATT 113
DB 1 ATGTCATCGAGAAATTTTAAGGTGAGTGTCTTAATGTGAATGACCGAGACTGAGATT 60
OY 114 CAGTCCGCTGTAACCTACGAAACCAACCGAATCTTGTTCAGAGAAACAGAAATGGCACTT 173
DB 61 CAGTCCGCTGTAACCTACGAAACCAACCGAATCTTGTTCAGAGAAACAGAAATGGCACTT 120
OY 174 CAGTGGATTTGTCAAAACCCAAATCTGTCAAAATGGAATTTAAACCAACATCCATGTTCT 233
DB 121 CAGTGGATTTGTCAAAACCCAAATCTGTCAAAATGGAATTTAAACCAACATCCATGTTCT 180
OY 234 AAATTAGGGGTAATGCTTGCGGTGGGGTGGAAACACGCGCTCAACCCCTCAACCGGTGAT 293
DB 181 AAATTAGGGGTAATGCTTGCGGTGGGGTGGAAACACGCGCTCAACCCCTCAACCGGTGAT 240
OY 294 GTTATTGCTAACCGAGAGGGCAATTCATGCGCTACAAAGGACAAAGATTCAACAGCCAT 353
DB 241 GTTATTGCTAACCGAGAGGGCAATTCATGCGCTACAAAGGACAAAGATTCAACAGCCAT 300
OY 354 TACCTTGGCTCCCTCAACCCCAACGCTCAGCTATCCGAGTTGGGTCTCCAGGAGAGGAA 413
DB 301 TACCTTGGCTCCCTCAACCCCAACGCTCAGCTATCCGAGTTGGGTCTCCAGGAGAGGAA 360
OY 414 ATCTATGCCCATTTCAAGAGCCGTCTTCCAAATGTTTAACTCGAGACATTTGTTTGGG 473
DB 361 ATCTATGCCCATTTCAAGAGCCGTCTTCCAAATGTTTAACTCGAGACATTTGTTTGGG 420
OY 474 GGATGGGATATCAACAACGTAACCTGCTGATGCCATGGCCAGAGCAAAAGTGTGATGAC 533
DB 421 GGATGGGATATCAACAACGTAACCTGCTGATGCCATGGCCAGAGCAAAAGTGTGATGAC 480
OY 534 ATCGATTTGCGAGAGAGGTTGAGGCGCTTACATGGAATCATGCTCCCGGAAATC 593
DB 481 ATCGATTTGCGAGAGAGGTTGAGGCGCTTACATGGAATCATGCTCCCGGAAATC 540
OY 594 TATGACCCCGGATTTTCAATGCTGCAACCAAGAGAGCGTGCACAAACGTCATCAAGGCG 653
DB 541 TATGACCCCGGATTTTCAATGCTGCTGCAACCAAGAGAGCGTGCACAAACGTCATCAAGGCG 600
OY 654 ACAAAAGCAAGACAGATTCAACAATCATCAAAAGACATCAAGGCGTTTAAAGAACCCACC 713
DB 601 ACAAAAGCAAGACAGATTCAACAATCATCAAAAGACATCAAGGCGTTTAAAGAACCCACC 660
OY 714 AAAGTGGACAAGGTGTTTACATGTTGAGTGCACACAGAGAGGTACGTAATTTGGTT 773
DB 661 AAAGTGGACAAGGTGTTTACATGTTGAGTGCACACAGAGAGGTATGCAATTTGGTT 720
OY 774 GTGAGGCTTTAATGACACATGAGAGATCTTTGGCTGCTGTGAGACAGAAATGAGGCTGAG 833
DB 721 GTGAGGCTTTAATGACACATGAGAGATCTTTGGCTGCTGTGAGACAGAAATGAGGCTGAG 780
OY 834 ATTTTCTCTTCCACTTGTATGCAATGCTTGTGTTATGAAAATGTTCTTTCAATTAAT 893
DB 781 ATTTTCTCTTCCACTTGTATGCAATGCTTGTGTTATGAAAATGTTCTTTCAATTAAT 840
OY 894 GGAAGCCCTCAAGACCTTTGTACAGAGGCTGATGATCTTGCCATGCGAGAGAACT 953
DB 841 GGAAGCCCTCAAGACCTTTGTACAGAGGCTGATGATCTTGCCATGCGAGAGAACT 900
OY 954 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGACCAAAATGAAATCTGTGTGTTGAT 1013
DB 901 TTGATTTGTGAGATGACTTCAAGAGTGTGACAGACCAAAATGAAATCTGTGTGTTGAT 960
OY 1014 TTCTTTGTGGGGGCTGGTATCAAGCCAACTTAATGTCAGTTTCAACCATCTGGGAAAC 1073

```

```

DB 961 TTTCTTGTGGGGGCTGGTATCAAGCCAAATCTATGTTAGTTACAAACCATCTGGGAAAC 1020
OY 1074 AATGATGATGAAATCTTTCGGCTCCACAAATCTTCGTTCCAAAGAAATCTCCAGAGC 1133
DB 1021 AATGATGATGAAATCTTTCGGCTCCACAAATCTTCGTTCCAAAGAAATCTCCAGAGC 1080
OY 1134 AACGTTGTTGATGATGTTCAACAGCAATGCCATCTCTTATGAGCCCTGTGTAACATCCA 1193
DB 1081 AACGTTGTTGATGATGTTCAACAGCAATGCCATCTCTTATGAGCCCTGTGTAACATCCA 1140
OY 1194 GACCAATGTTGTTATTAAGTATGTCCTTACGTAAGGAGGACAGCAAGAGCCATGAT 1253
DB 1141 GACCAATGTTGTTATTAAGTATGTCCTTACGTAAGGAGGATAGCAAGAGCCATGAT 1200
OY 1254 GAGTACACTTCAGAGATATTCATGAGTGGAAAGACACATTTGTTTTCACAAACATGTC 1313
DB 1201 GAGTACACTTCAGAGATATTCATGAGTGGAAAGACACATTTGTTTTCACAAACATGTC 1260
OY 1314 GAGGATTCCTCTTATGCTGCTCTTATTTATCTTGAATGTTGTTCTTCTGAGCTGAGC 1373
DB 1261 GAGGATTCCTCTTATGCTGCTCTTATTTATCTTGAATGTTGTTCTTCTGAGCTGAGC 1320
OY 1374 ACTAGAAATCGAGTTTAAAGTGAAGGAAATGAGGAAATTCATCTCATTCACCAATGCT 1433
DB 1321 ACTAGAAATCGAGTTTAAAGTGAAGGAAATGAGGAAATTCATCTCATTCACCAATGCT 1380
OY 1434 ACCATCTCAGCTACCTCAGCAAGGCTCTCTGTTCCACCGGGTACCAAGTGTGAT 1493
DB 1381 ACCATCTCAGCTACCTCAGCAAGGCTCTCTGTTCCACCGGGTACCAAGTGTGAT 1440
OY 1494 GCATTGTCAAAAGCAGCGTGCATGCTGGAATAATGAGGGCTTGTGTTGATGGCC 1553
DB 1441 GCATTGTCAAAAGCAGCGTGCATGCTGGAATAATGAGGGCTTGTGTTGATGGCC 1500
OY 1554 CCAGGAATTAACATGATCTCGAGTACAAATGTA 1586
DB 1501 CCAGGAATTAACATGATCTCGAGTACAAATGTA 1533

RESULT 13
ADQ14500
ID ADQ14500 standard; cDNA, 1533 BP.
XX
XX ADQ14500;
XX
XX 23-SEP-2004 (first entry)
XX
DE Mutant soybean myo-inositol 1-phosphate synthase cDNA #3.
XX
XX Soybean; myo-inositol 1-phosphate synthase; gene; ss;
KW myo-inositol 1-phosphate; raffinose; stachyose; phytic acid; sucrose;
KW inorganic phosphate; mutant.
XX
OS Glycine max.
OS Synthetic.
OS
XX
XX Key Location/Qualifiers
XX CDS 1..1533
XX FT /product= "Mutant soybean myo-inositol 1-phosphate
XX FT /tag= a
XX FT /product= #3"
XX
XX US2004128713-A1.
XX
XX 01-JUL-2004.
XX
XX 21-NOV-2003; 2003US-00718952.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX 26-APR-1999; 98US-00293315.
XX 11-MAR-2002; 2002US-00025003.
XX

```


RESULT 14
ADS82001
ID ADS82001 standard; cDNA; 1533 BP.
XX ADS82001;
AC
DT 18-NOV-2004 (first entry)
XX
DE Soybean myo-inositol 1-phosphate synthase cDNA mutant #2.
XX
KM Soybean; plant; myo-inositol 1-phosphate synthase; ss; gene; phytic acid;
XX raffinose; stachyose; sucrose; inorganic phosphate; flatulence; mutant.
XX
OS Glycine max; line 29010CP01.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1533
FT /tag= a
FT /product= "myo-inositol 1-phosphate synthase"
FT mutation /replace(260,G)
FT /tag= b
XX
XX US2003074685-A1.
XX
XX 17-APR-2003.
XX
XX 11-MAR-2002; 2002US-00025003.
XX
XX 08-APR-1997; 97US-00835751.
XX 07-APR-1998; 98WO-US006822.
XX
XX (HITZ/) HITZ W D.
XX (SEBA/) SEBASTIAN S A.
XX
XX Hitz WD, Sebastian SA;
XX
XX WPI: 2004-639957/62.
XX P-PSDB; ADS82002.
XX
XX Novel isolated nucleic acid fragment encoding soybean myo-inositol 1-phosphate synthase, useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds.
XX
XX Claim 8; SEQ ID NO 11; 34pp; English.
XX
XX The invention relates to an isolated nucleic acid fragment encoding a soybean myo-inositol 1-phosphate synthase, or a mutant myo-inositol 1-phosphate synthase having decreasing capacity for the synthesis for myo-inositol-1-phosphate. Also included are a chimeric gene (comprising the nucleic acid encoding soybean myo-inositol 1-phosphate synthase, its complement, subfragment or the complement of the subfragment, operably linked to suitable regulatory sequences, where expression of the chimeric gene results in a decrease in expression of an endogenous or native gene encoding a soybean myo-inositol 1-phosphate synthase), a soybean plant comprising the chimeric gene (with a heritable phenotype of a seed phytic acid content of less than 17 micromol/g, a seed content of raffinose plus stachyose of less than 14.5 micromol/g, and a seed sucrose content of greater than 200 micromol/g, provided that the plant is not LR33), seeds from the plant, making a soybean plant with the heritable phenotype (comprising crossing LR33 or the plant comprising the chimeric gene with an elite soybean plant and selecting a progeny plant of the cross of crossing step that has a heritable phenotype as mentioned above), seeds of soybean plant made by the above method, a soy protein product derived from seeds of a soybean plant (homozygous for one or more gene encoding a mutant myo-inositol 1-phosphate synthase having decreased capacity for the synthesis of myo-inositol 1-phosphate, where the gene confers a heritable phenotype as mentioned above), and making or producing a soybean protein product derived from seeds of a soybean plant with heritable phenotype as mentioned above. The nucleic acid is useful for altering raffinose saccharide, sucrose, phytic acid and inorganic phosphate content of soybean seeds thus leading to valuable and useful

CC soybean products, since the presence of high concentration of raffinose
CC oligosaccharides in soy plants (and other legumes) can lead to flatulence
CC when consumed by humans. The present sequence encodes a mutant myo-
CC inositol 1-phosphate synthase.
XX

SQ Sequence 1533 BP, 430 A; 341 C; 370 G; 392 T; 0 U; 0 Other;

Query Match 83.2%; Score 1464.2; DB 13; Length 1533;

Best Local Similarity 97.2%; Pred. No. 0;

Matches 1490; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 54 ATGTTTCATCGAATTTTAAAGTTGAGTGTCTTAATGGAAGTACACGAGACTGAGATT 113
DB 1 ATGTTTCATCGAATTTTAAAGTTGAGTGTCTTAATGGAAGTACACGAGACTGAGATT 60
QY 114 CAGTCCGTTGTCACAACTAGCAAAACCCGAACTGTTTCAGAGAACAGAAATGACCTAT 173
DB 61 CAGTCCGTTGTCACAACTAGCAAAACCCGAACTGTTTCAGAGAACAGAAATGACCTAT 120
QY 174 CAGTGGATTGTCAAACCCGAACTGTTTCAGAAATGCAATTTAAACCAACATCCATGTTCT 233
DB 121 CAGTGGATTGTCAAACCCGAACTGTTTCAGAAATGCAATTTAAACCAACATCCATGTTCA 180
QY 234 AAATTAGGGGTATATGCTTGTGGGTGGGTGGAACAAACGCTCAACCTCACCCTGTGT 293
DB 181 AAATTAGGGGTATATGCTTGTGGGTGGGTGGAACAAACGCTCACCCTCACCCTGTGT 240
QY 294 GTTATTGTCACCCGAGAGGGGCTTTCATGGGCTACAAAGACAAAGATTCAACAAACCAAT 353
DB 241 GTTATTGTCACCCGAGAGGGGCTTTCATGGGCTACAAAGACAAAGATTCAACAAACCAAT 300
QY 354 TACTTTGGCTCCCTCACCAGGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGAA 413
DB 301 TACTTTGGCTCCCTCACCAGGCTCAGCTATCCGAGTTGGGTCTTCCAGGAGAGAA 360
QY 414 ATCTATGCCCATTTAAGAGCTGCTTCCAAATGTTAACCTTGACGACAAATGTTTGGG 473
DB 361 ATCTATGCCCATTTAAGAGCTGCTTCCAAATGTTAACCTTGACGACAAATGTTTGGG 420
QY 474 GGATGGATATCAGCAATGAACTGCTGATGCTATGCTGACGAGGCAAGAGTGTGTGAC 533
DB 421 GGATGGATATCAGCAATGAACTGCTGATGCTATGCTGACGAGGCAAGAGTGTGTGAC 480
QY 534 ATCGATTTGCAAGAACAGTTGAGGCTTACATGAAATCATGATCTTCCATCCCGGAATC 593
DB 481 ATCGATTTGCAAGAACAGTTGAGGCTTACATGAAATCATGATCTTCCATCCCGGAATC 540
QY 594 TATGACCCGGATTTTATTTGCTGCTCCAAACCAAGAGGCTGCCAACAAGTATCAAGGGC 653
DB 541 TACGACCCGGATTTTATTTGCTGCTCCAAACCAAGAGGCTGCCAACAAGTATCAAGGGC 600
QY 654 ACAAGCAAGAGCAAGTTCAACAAATCATCAAGACATCAAGGCTTTAAGAACCCACC 713
DB 601 ACAAGCAAGAGCAAGTTCAACAAATCATCAAGACATCAAGGCTTTAAGAACCCACC 660
QY 714 AAAGTGCACAAAGTGTGTGATCTGTGACCTGCAACACAGAGAGGTACAGTAATTTGGTT 773
DB 661 AAAGTGCACAAAGTGTGTGATCTGTGACCTGCAACACAGAGAGGTATAGCAATTTGGTT 720
QY 774 GTGGGCTTTAATGACCAATGAGAAATCTTTGCTGCTGTGTGACAGAAATGAGGCTGAG 833
DB 721 GTAGGCTTTAATGACCAATGAGAAATCTTTGCTGCTGTGTGACAGAAATGAGGCTGAG 780
QY 834 ATTTCTCCCTGACCTTGTATGAGCCATGCTGTGTGTATGAAATAGTTCCTTCAATTA 893
DB 781 ATTTCTCCCTGACCTTGTATGAGCCATGCTGTGTGTATGAAATAGTTCCTTCAATTA 840
QY 894 GGAAGCCCTCAGAACACTTTTGTACACAGGCTGATTTGATCTTCCATGCGAGAACACT 953
DB 841 GGAAGCCCTCAGAACACTTTTGTACACAGGCTGATTTGATCTTCCATGCGAGAACACT 900
QY 954 TTGATTTGGTGAATGACTTCAAGAGTGTGACCAACAAATGAATCTGTGTGTGAT 1013

DB 901 TTGATGTGTGAGATGACTTCAAGAGTGGTCAGACCAAAATGAAATCTGNTGTTGAT 960
QY 1014 TTCTTTGGGGGGCTGTATCAAGCCAACTATATGTAGTTACAAACCTCTGGGAAAC 1073
DB 961 TTTCTTGTGGGGCTGTATCAAGCCAACTATATGTAGTTACAAACCTCTGGGAAAC 1020
QY 1074 AATGATGTATGATCTTTCGGCTCCAAACCTTTCGGTCCAAAGAAATCTCCAGAGC 1133
DB 1021 AATGATGTATGATCTTTCGGCTCCAAACCTTTCGGTCCAAAGAAATCTCCAGAGC 1080
QY 1134 AACGTTGTATGATATGTGTCAACAGCAATGCCATCTCTATAGAGCTGTGAACATCA 1193
DB 1081 AACGTTGTATGATATGTGTCAACAGCAATGCCATCTCTATAGAGCTGTGAACATCC 1140
QY 1194 GACCATGTGTGTATTTAAGTATGTCTTACGTAAGGGAGACAGAAAGAGCCATGAT 1253
DB 1141 GACCATGTGTGTATTTAAGTATGTCTTACGTAAGGGAGATAGCAAGAGCCATGAT 1200
QY 1254 GAGTACACTTCAGAGATATTCATGGGTGAAAGAGACATTTGTTGACAAACATATGC 1313
DB 1201 GAGTACACTTCAGAGATATTCATGGGTGAAAGAGACATTTGTTGACAAACATATGC 1260
QY 1314 GAGATTCCTCTGTAGCTCTCTATATCTTGAATCTTGTCTGTGAGCTCAGC 1373
DB 1261 GAGATTCCTCTGTAGCTCTCTATATCTTGAATCTTGTCTGTGAGCTCAGC 1320
QY 1374 ACTAGATGAGTTTAAAGCTGAAATGAGGAAATTCACATCTTCCACCCAGTTGCT 1433
DB 1321 ACTAGATGAGTTTAAAGCTGAAATGAGGAAATTCACATCTTCCACCCAGTTGCT 1380
QY 1434 ACCATCTCAGCTACCTCAACAAAGCTCTCTGTTCCACCGGTACACAGTGTGAT 1493
DB 1381 ACCATCTCAGCTACCTCAACAAAGCTCTCTGTTCCACCGGTACACAGTGTGAT 1440
QY 1494 GCATTGTCAAGCAGCTGCAATGTGAAATCATATAGAGGCTTGTGTGATGGCC 1553
DB 1441 GCATTGTCAAGCAGCTGCAATGTGAAATCATATAGAGGCTTGTGTGATGGCC 1500
QY 1554 CCAGAGATATACATGATTTCTCGATACAAAGTGA 1586
DB 1501 CCAGAGATATACATGATTTCTCGATACAAAGTGA 1533

RESULT 15
AAK90402
ID AAK90402 standard; cDNA to mRNA; 1950 BP.
AC AAK90402;
XX
XX 24-SEP-1999 (first entry)
DB Nicotiana paniculata INPS encoding cDNA.
XX Nicotiana paniculata; INPS; NpINPS1; inositol monophosphate synthase;
KW water stress; resistance; de.
XX Nicotiana paniculata.
OS
XX
XX JPI1187879-A.
PN
XX
XX 13-JUL-1999.
PD
XX
XX 26-DEC-1997; 97JP-00359773.
PF
XX
XX 26-DEC-1997; 97JP-00359773.
PR
XX
XX (NISB) JAPAN TOBACCO INC.
PA
XX
XX WPI: 1999-451546/38.
DR P-PSDB; AAY24477.
XX
XX New INPS gene derived from Nicotiana genus plant - useful for conferring
PT resistance to water stress to plants.

XX
PS Claim 2; Page 6-8; Bpp; Japanese.
XX
CC The present sequence encodes Nicotiana paniculata inositol monophosphate
CC synthase (INPS), designated NpINPS1. INPS can be used to confer water
CC stress resistance to a plant
XX
SQ Sequence 1950 BP; 570 A; 377 C; 425 G; 578 T; 0 U; 0 Other;
Query Match 63.7%; Score 1120.4; DB 2; Length 1950;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 1277; Conservative 0; Mismatches 261; Indels 0; Gaps 0;
QY 51 ATTAATGTTATGAGATTTTAAAGTTGATGCTCTTAATGTGAATGACCCGAGCTGAG 110
DB 89 AAAATGTTATGAGATTTTAAAGTTGATGACCCCGAGCTTGAATGACCCGAGCTGAG 148
QY 111 ATTCAAGTCCGTGTACACTACGAAACCAACGAACTGTTACAGAGAACAGATGACAC 170
DB 149 ATTCACTCTGTCTATGATTTATCAAAACCACTGAGTTATGATGAGAAATGAGACA 208
QY 171 TATCAGTGTATGTCAAAACCAATCTGTCAAAATGCAATTTAAACCAACATCCATGTT 230
DB 209 TATCAATGAGACCGTCAAGCTTAAGACTGTCAAAATGAGTTCAAGCTGATGTTATGTT 268
QY 231 CCTAAATTAAGGGTAAATGCTTGTGGTGGGTGAAACACAGGCTCAACCTCACCCT 290
DB 269 CCCAAATTAAGGGGTAAATGCTTGTGGTGGGTGAAACCAATGTTCAACCTTGACCCGT 328
QY 291 GGTGTATATGCTAACCGAGAGGGCAATTCATGGGGCTACAAAGAGACAAAGATTCACAGCC 350
DB 329 GGTGTATATGCTAACCGAGAGAAATTCATGGGGCTACATGGGCAACCAAGATTAAGTGCACAGCC 388
QY 351 AATTAATTTGGCTCCCTCAACCAAGCTCAGCTATCCAGTTGGGTCTTTCAGAGGAGAG 410
DB 389 AATTAATTTGGCTCTTCACTCAAGGCTTCACTATTCAGAGTTGGGTCTTTCAGATGAGAA 448
QY 411 GAAATCTATGCCCCCATTCAGAGCTGCTTCCAAATGTTTAAACCTGACCAATGTGTTT 470
DB 449 GAAATCTATGCTCCCATTTAAAGCTCTCTTCCAAATGTTTAAACCTGACCAATGTGTTT 508
QY 471 GGGGGATGGGATATAGCAACAGTAAACCTGCTGATGCTGATGCTGATGCTGATGCTGAT 530
DB 509 GGAAGATGGGATATAGCAACAGTAAACCTGCTGATGCTGATGCTGATGCTGATGCTGAT 568
QY 531 GACATGATTTGACAGAGCAGTTGAGGCTTTCATGATGATCCATGCTTCCATCCCGGA 590
DB 569 GATATGATTTGACAGAGCAGTTGAGGCTTTCATGATGATCCATGCTTCCATCCCGGT 628
QY 591 ATCTATGACCGGATTTTCATGCTGCTCCCAACAAAGAGAGCTGCCAACACGTCATCAG 650
DB 629 ATCTATGACCGGATTTTCATGCTGCTCCCAACAAAGAGAGCTGCCAACACGTCATCAG 688
QY 651 GGCACAAAGCAAGACCAATTCACAAATCATCAAGATCAATCAAGGCTTTAAGAGAGCC 710
DB 689 GGAACCAAGAAAGCAAAATGATCAAAATCATTAAGATTAAGAGAGTTAAGAGAGAG 748
QY 711 ACCAAAGTGAACAAGTGTGTGATCTGATGATGCTGCAACACAGAGAGTACAGTAAATTTG 770
DB 749 AACAAAGTGAACAAGTGTGTGATCTGATGATGCTGCAACACAGAGAGTACAGTAAATTTG 808
QY 771 GTTGTGGGCTTTAATGACCACTGAGAAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 830
DB 809 GTTGTGTGCTTTAATGACCACTGAGAAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 868
QY 831 GAGATTTCTCTTCAACCTGATGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 890
DB 869 GAAATATCTCTTCAACCTGATGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928
QY 891 AATGAGCCCTCAGAACATTTTGTACAGGCTGATGATCTTGCATTCGACGAGAAC 950
DB 929 AATGAGCCCTCAGAACATTTTGTACAGGCTGATGATCTTGCATTCGACGAGAAC 988

```
OY 951 ACTTGATTGGTGAGATGACTTCAAGATGCTCAGACCACCAAAATGAAATCTGTGGTT 1010
    |||||
Db 989 ACATTGATTGGTGGTGAAGTACTTTAAGATGCTCAACCAAAATGAAGTCACTGGTT 1048
    |||||
OY 1011 GATTTCTTTGGGGGCTGGTATCAAGCCACATCTAATAGTCAATTACACCATCTGGGA 1070
    |||||
Db 1049 GATTTCTTTGGTGGAGCTGGTATTAAGCCACATCAATGTGAGCTACACCAATTTGGGT 1108
    |||||
OY 1071 AACATGATGATGATGAATTTCCGGCTCCAGAACTTCCGTTCCAGGAATCTCCAAG 1130
    |||||
Db 1109 AACAAATGATGATGAATTTCTGCTCCCTCAACTTCCGGTCAAGGAGATCTGAAA 1168
    |||||
OY 1131 AGCAAGTTGTGATGATGATGTGTCAAAGCAATGCCATCCTATAGAGCTGTGAACAT 1190
    |||||
Db 1169 AGTAATGTTGTGATGATGATGTTTCAAGCAATGCCATCCTTATAGAGCTGTGAAGCAC 1228
    |||||
OY 1191 CCAGACCATGTTGTTTATTAAGTATGTGCTTACGTAGGGGACAGCAAGAGCCATG 1250
    |||||
Db 1229 CCTGACCATGTTGTTGATTAAGTATGTGCCATATGTGGAGACAGCAAGAGGCCAATG 1288
    |||||
OY 1251 GATGAGTACACTTCAGAGATATTCATGGGTGGAAGAGCACCATGTTTGCACAAACA 1310
    |||||
Db 1289 GATGAGTACACATCTGAGATTTTCAATGGGGGAAAGAACACCATGTTTGCACAAATCT 1348
    |||||
OY 1311 TSCGAGATTCCCTCTTAAGCTGCTCCTATATCTTGACTGTGCTTCTTGCTGAGCTC 1370
    |||||
Db 1349 TGTGAGATTCACTTTAGCTGCTCCAATTATATTGGAATTGTGCTTCTTGCTGAGCTC 1408
    |||||
OY 1371 AGCACTAGATGAGTTTAAAGCTGAAAATGAGGGAAAATTCCACTCATTCACCCAGTT 1430
    |||||
Db 1409 AGTACCCGCAATCAGCTCAAGCTGAAGGAGAGGGTAAAGTCCACTCTTCCACCCCGTG 1468
    |||||
OY 1431 GGTACCATCTCAGCTACTCAACCAAGGCTCCTGTGTTCCACCGGTTACACCAATGTTG 1490
    |||||
Db 1469 GCTACTATCTCAGCTAACCTTACCAAGGCTCCTGTGTTACACCAAGTTACCAAGTGTG 1528
    |||||
OY 1491 AATGCAATGTCAAAGCAGCGTGAATGCTGAAAAATATATGAGGGCTGTGTGGATTG 1550
    |||||
Db 1529 AATGCACTCTCAAGGAGAGGGCAATGCTTGAGAACTATATGAGGGCTGTGTGGACTT 1588
    |||||
OY 1551 GCCCCAGAGATTAACATGATTTCTCGAGTACAAAGTGAAG 1588
    |||||
Db 1589 GCACCCAGAGAACAAATGATTTCTGGAATACAAATGAAG 1626
    |||||
```

Search completed: June 7, 2005, 22:54:21
Job time : 981.454 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2005, 20:51:42 ; Search time 7677.08 Seconds
(without alignments)
1108.537 Million cell updates/sec

Title: US-10-718-952-1

Perfect score: 1760
Sequence: 1 cctctctatctcctcttctgt.....aatgtagtataatttctgt 1760

Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapect 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenBank1:*

- 1: gb_ba:*
- 2: gb_hcg:*
- 3: gb_in:*
- 4: gb_cm:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sta:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1760	100.0	1782	6	BD075266 Soybean p
2	1750.4	99.5	1781	8	AF293970 Glycine m
3	1717	97.6	1739	8	AY038802 Glycine m
4	1531.4	87.0	1533	6	BD075269 Soybean p
5	1120.4	63.7	1931	8	AB032073 Nicotiana
6	1120.4	63.7	1931	8	AB032073 Nicotiana
7	1115.6	63.4	1950	6	E27176 Novel INPS
8	1092.2	62.1	1954	8	AB009881 Nicotiana
9	1087	61.8	1959	8	BT013759 Lycopersi
10	1073.4	61.0	1978	8	BT013505 Lycopersi
11	1064.2	60.5	1845	8	AF284065 Sesamum i
12	1037.8	59.0	2053	8	MCU32511 Mesembryant
13	1037.4	58.9	1986	8	AY028259 Aricennia
14	1010.8	57.4	1863	8	AF433879 Sauea ma
15	1008.4	57.3	1863	8	AY065415 Arabidops
16	1006.8	57.2	1863	8	AY065415 Arabidops
17	1006.8	57.2	1863	8	AY054202 Arabidops
18	1006.8	57.2	1864	8	AY053415 Arabidops
19	1002	56.9	1890	8	ATU30250 Arabidops

20	1001.4	56.9	1781	6	AX054630 Sequence
21	1001.4	56.9	1781	8	BN06307 Brassica na
22	999.6	56.8	1538	8	AF120146 Trifolium
23	999.6	56.8	1538	8	AF120147 Trifolium
24	999.6	56.8	1538	8	AF120148 Trifolium
25	999.2	56.8	1914	8	AK103501 Oryza sat
26	999.2	56.8	1915	8	AK058750 Oryza sat
27	997.6	56.7	1868	8	AB012107 Oryza sat
28	996.2	56.6	1870	8	AF372954 Arabidops
29	995.4	56.6	1533	6	CQ0805008 Sequence
30	995.4	56.6	1533	6	AX0506743 Sequence
31	995.4	56.6	1533	8	AY143904 Arabidops
32	995.4	56.6	1533	8	AY143904 Arabidops
33	994.6	56.5	1837	8	AY085931 Arabidops
34	994.2	56.5	1959	6	AX151207 Sequence
35	990.8	56.3	1567	8	BT001931 Arabidops
36	990.8	56.3	1665	6	BD073470 Regulated
37	990.8	56.3	1665	6	AF056326 Zea mays
38	990.4	56.3	1921	8	ATU04876 Arabidops
39	988.8	56.2	1871	8	PVU38920 Phaseolus v
40	987.4	56.1	1535	6	BD073472 Regulated
41	985.6	56.0	2152	8	AF056325 Hordeum v
42	984.6	55.9	1936	8	AB059557 Avena sat
43	984	55.9	1931	6	AR137808 Sequence
44	984	55.9	1931	6	AR170167 Sequence
45	982.2	55.8	1930	8	AY035045 Arabidops

ALIGNMENTS

RESULT 1	BD075266	1782 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD075266				
DEFINITION	Soybean plant producing seeds with reduced levels of raffinose				
ACCESSION	BD075266.1	GI:22620869			
VERSION	JP 2001519665-A/1.				
KEYWORDS	unidentified				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1782)				
AUTHORS	Hitz,W.D. and Sebastian,S.A.				
TITLE	Soybean plant producing seeds with reduced levels of raffinose				
JOURNAL	saccharides and phytic acid				
COMMENT	Patent: JP 2001519665-A 1 23-OCT-2001;				
	SI DU PONT DE NEMOURS AND CO				
	OS Soybean line LR13				
	PN JP 2001519665-A/1				
	PD 23-OCT-2001				
	PF 07-APR-1998 JP 1998543012				
	PR 08-APR-1997 US 08/835751				
	PI WILLIAM DEAN HITZ, SCOTT ANTHONY SEBASTIAN				
	PC C12N15/52, C12N15/82, C12N15/11, C12N9/90, A01H5/00 CC				
	Strandedness: Double;				
	CC Topology: Linear;				
	CC Soybean plant producing seeds with reduced				
	levels of raffinose				
	CC and phytic acid				
	FM Key Location/Qualifiers				
	FT CDS 54..1586.				
FEATURES	source				
	1..1782				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
ORIGIN	Query Match	100.0%; Score 1760;	DB 6;	Length 1782;	
	Best Local Similarity	100.0%; Pred. No. 0;			
	Matches 1760;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

OY 1 CTTCTTTATTCCTTTTGTATTCATTCATCTTAACTTTGTGAAAAATATGTTCA 60
DB 1 CTTCTTTATTCCTTTTGTATTCATTCATCTTAACTTTGTGAAAAATATGTTCA 60
OY 61 TCGAGAAATTTAAGGTGAGTGTCTTAATGTAGTACCGGAGCTGATTCAGTCCG 120
DB 61 TCGAGAAATTTAAGGTGAGTGTCTTAATGTAGTACCGGAGCTGATTCAGTCCG 120
OY 121 TGTACAACTACGAAACCAACCGAATCTTTCACGAGAACGAAATGCACTATCAGTGA 180
DB 121 TGTACAACTACGAAACCAACCGAATCTTTCACGAGAACGAAATGCACTATCAGTGA 180
OY 181 TTGTCAAAACCAATCTGCAATACGAATTTAAAAACCAACCAATCTTCTTAATTAG 240
DB 181 TTGTCAAAACCAATCTGCAATACGAATTTAAAAACCAACCAATCTTCTTAATTAG 240
OY 241 GGGTAATGCTTGTGGGTGGGGTGGAAACACGGCTCAACCCCTCACCGGTGTGTTATTG 300
DB 241 GGGTAATGCTTGTGGGTGGGGTGGAAACACGGCTCAACCCCTCACCGGTGTGTTATTG 300
OY 301 CTAAACGAGAGGGGCAATTTATGGGCTACAAAGACAAAGTTCAACAGCCATTTACTTTG 360
DB 301 CTAAACGAGAGGGGCAATTTATGGGCTACAAAGACAAAGTTCAACAGCCATTTACTTTG 360
OY 361 GCTCCCTCACCAAGGCTCAGCTATCCGAGTTGGGTCTTCCAGGGAGAGAAATCTATG 420
DB 361 GCTCCCTCACCAAGGCTCAGCTATCCGAGTTGGGTCTTCCAGGGAGAGAAATCTATG 420
OY 421 CCCCATTCAGAGCCTGCTTCCAAATGTTAACCTGTGACGACATTTGTTGGGGAGATGG 480
DB 421 CCCCATTCAGAGCCTGCTTCCAAATGTTAACCTGTGACGACATTTGTTGGGGAGATGG 480
OY 481 AATATGACAACTGAAACCTGCGTGAATGCCATGGCCAGGGCAAGGTGATGATCGAT 540
DB 481 AATATGACAACTGAAACCTGCGTGAATGCCATGGCCAGGGCAAGGTGATGATCGAT 540
OY 541 TGCAGAGAGAGTGGAGCCTTACATGAGATTCATGCTTCCCGGCAATCTATGACC 600
DB 541 TGCAGAGAGAGTGGAGCCTTACATGAGATTCATGCTTCCCGGCAATCTATGACC 600
OY 601 CGGATTTCTTGTCTGCAACCAAGAGAGCGTGCACCAACGTCATCAAGGGCACAAGC 660
DB 601 CGGATTTCTTGTCTGCAACCAAGAGAGCGTGCACCAACGTCATCAAGGGCACAAGC 660
OY 661 AAGAGCAAGTTCAAAATCATCAAGAGATCAAGCCGTTAAGAGGACACCAAGTGG 720
DB 661 AAGAGCAAGTTCAAAATCATCAAGAGATCAAGCCGTTAAGAGGACACCAAGTGG 720
OY 721 ACAAGGTGTTGATCTGTGAGCTGCAACAGAGAGTACAGTAAATTTGTTGGGCC 780
DB 721 ACAAGGTGTTGATCTGTGAGCTGCAACAGAGAGTACAGTAAATTTGTTGGGCC 780
OY 781 TTAATGACACCAATGAGAAATCTTGTGGCTGTGTGACAGAAATGAGGCTGATTTCTC 840
DB 781 TTAATGACACCAATGAGAAATCTTGTGGCTGTGTGACAGAAATGAGGCTGATTTCTC 840
OY 841 CTTCCACTTTGATGTCATTTGCTGTGTTATGAAAAATTTCTTCAATTAATGAAAGC 900
DB 841 CTTCCACTTTGATGTCATTTGCTGTGTTATGAAAAATTTCTTCAATTAATGAAAGC 900
OY 901 CTCAGAACTTTTGTACAGGGCTGATGATCTTCCATCGGAGAGACATTTGATTTG 960
DB 901 CTCAGAACTTTTGTACAGGGCTGATGATCTTCCATCGGAGAGACATTTGATTTG 960
OY 961 GTGAGATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGTTGATTTCTTG 1020
DB 961 GTGAGATGATCTTCAAGAGTGTGACAGCAAAATGAAATCTGTGTTGTTGATTTCTTG 1020
OY 1021 TGGGGGCTGTATCAAGCCAACTCTATAGTCAAGTTACAACTCTGGGAAACAATGATG 1080
DB 1021 TGGGGGCTGTATCAAGCCAACTCTATAGTCAAGTTACAACTCTGGGAAACAATGATG 1080

```

```

OY 1081 GTATGAATCTTTCGGCTCCACAAACCTTTCGGTTCAGAGAAATCTCCAGAGCAACGTTG 1140
DB 1081 GTATGAATCTTTCGGCTCCACAAACCTTTCGGTTCAGAGAAATCTCCAGAGCAACGTTG 1140
OY 1141 TTGATGATATGATGTCACACGCAATGTCATCTCTATGAGCTGTGTGAATATCCAGCCATG 1200
DB 1141 TTGATGATATGATGTCACACGCAATGTCATCTCTATGAGCTGTGTGAATATCCAGCCATG 1200
OY 1201 TTGTTGTTAATTAAGATGTGCTTACGTAGGGGACAGAGAGAGCCATGATAGTAC 1260
DB 1201 TTGTTGTTAATTAAGATGTGCTTACGTAGGGGACAGAGAGAGCCATGATAGTAC 1260
OY 1261 CTTGAGATATTCATGAGGTGGAAAGAGACACATTTTTCACAAACATGCGAGATT 1320
DB 1261 CTTGAGATATTCATGAGGTGGAAAGAGACACATTTTTCACAAACATGCGAGATT 1320
OY 1321 CCCTCTTAGCTGCTCTTATTTATCTTGGACCTTGCTTCTTGTGAGCTCAGCACTAAG 1380
DB 1321 CCCTCTTAGCTGCTCTTATTTATCTTGGACCTTGCTTCTTGTGAGCTCAGCACTAAG 1380
OY 1381 TCGAGTTTAAAGCTGAAATGAGGAAATTCACATCTCATTCACCCAGTGTCTACATCC 1440
DB 1381 TCGAGTTTAAAGCTGAAATGAGGAAATTCACATCTCATTCACCCAGTGTCTACATCC 1440
OY 1441 TCAAGTACCTCAACCAAGGCTCCTGTGTTCACCGGGTACACAGTGTGAATGATTTGT 1500
DB 1441 TCAAGTACCTCAACCAAGGCTCCTGTGTTCACCGGGTACACAGTGTGAATGATTTGT 1500
OY 1501 CAAGACACGTCGAATGCTGTGAAAAATTAATGAGGCTTGTGTTGATTTGGCCCGAGAG 1560
DB 1501 CAAGACACGTCGAATGCTGTGAAAAATTAATGAGGCTTGTGTTGATTTGGCCCGAGAG 1560
OY 1561 ATTAATGATTTCTCGAGTACAAAGTAAAGTAAAGTAAAGTAAATTAATGTTGGGGTAG 1620
DB 1561 ATTAATGATTTCTCGAGTACAAAGTAAAGTAAAGTAAAGTAAATTAATGTTGGGGTAG 1620
OY 1621 CCTAGCTGAATGTTTATGTTAATTAATTAATGTTTCTTAAATTTTGAAGTAAATTTGAA 1680
DB 1621 CCTAGCTGAATGTTTATGTTAATTAATTAATGTTTCTTAAATTTTGAAGTAAATTTGAA 1680
OY 1681 TGCATCAGCTTCATTAATGCTTTAGAGGGGCAATTCGTGTTACTGAGAAACATGAATG 1740
DB 1681 TGCATCAGCTTCATTAATGCTTTAGAGGGGCAATTCGTGTTACTGAGAAACATGAATG 1740
OY 1741 AATGATATTAATTTTGTGT 1760
DB 1741 AATGATATTAATTTTGTGT 1760

```

RESULT 2
 AF293970 1791 bp mRNA linear PLN 03-MAY-2001
 LOCUS Glycine max myo-inositol-3-phosphate synthase (MIRS) cDNA, complete
 DEFINITION
 cdb.
 ACCESSION AF293970
 VERSION AF293970.1 GI:13936690
 KEYWORDS
 SOURCE
 ORGANISM
 Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 1791)
 REFERENCE
 1 Hegeman, C.E., Good, L.L. and Grabau, E.A.
 Expression of D-myo-inositol-3-phosphate synthase in soybean.
 Implications for phytic acid biosynthesis
 Plant Physiol. 125 (4), 1941-1948 (2001)
 JOURNAL MEDLINE
 PUBMED 21196082
 11293373
 2 (bases 1 to 1791)
 REFERENCE
 2 Hegeman, C.E., Good, L.L. and Grabau, E.A.
 Direct Submission

JOURNAL Submitted (07-AUG-2000) Plant Pathology, Physiology, and Weed Science, Virginia Tech, Fralin Biotechnology Center, Blacksburg, VA 24060, USA

FEATURES
source Location/Qualifiers

1..1791
/organism="glycine max"
/mol_type="mRNA"
/cultivar="William 82"
/db_xref="taxon:3847"
/dev_stage="Immature cotyledons"
1..1791
/gene="MIPS"
62..1594
/gene="MIPS"
/note="isomerase"
/codon_start=1
/product="myo-inositol-3-phosphate synthase"
/protein_id="AAK49896.1"
/db_xref="GI:13936691"
/translation="MPIENPKYECBPKVKTETELIOSVYVETTELVHNRNGYQWY
KKSXYREKTNHVPKLGMLVGMGNNGSTLGGVIANRBSISATYDKIQANVF
GSLTQASAIRVGSFQEEIYAPPKSLIPVNPDDIVFGWDISNMNLADAMARAFD
IDLOQMRPYMESMLPLPGIYDPDFIAQOEBRANNVIGTQEQVQQIIXDIKAKE
ATKVDKVVVLTANTERYSNLVGLNDTMENTLAAVDREABISPTLYAICVENV
PINGSPOITFVPLGLIDLAIRNTLIGDDPKSGTKMSVLDPLVGAGIKPTSVS
VNHGNNDDGMNLSAQOTPRSKISKSNVYDMMVNSAIIYEGEHPDHVYIKYPIY
GDSKAMDEYTSIEIFMGKSTIVLHNTCDLSLAETIIDLVLAEVLSTRIFKANE
GKFSPHFPVATLLSYLTAKPLVPGPTGVNALSKQAMLENIMRAICVGLAPENMILE
YK"

ORIGIN

Query Match 99.5%; Score 1750.4; DB 8; Length 1791;
Beet Local Similarity 99.7%; Pred. No. 0;

Matches 1754; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
1 CTCTCTTATTCCTTTGTAATTTGATTCATTCCTTAATCTTTGTGAAAAATATGTCA 60
9 CTCTCTTATTCCTTTGTAATTTGATTCATTCCTTAATCTTTGTGAAAAATATGTCA 68
61 TCGAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCGAGACTGAGATTCACTCG 120
69 TCGAGATTTTAAAGTTGAGTGTCTTAATGTGAAGTACCGAGACTGAGATTCACTCG 128
121 TGTACAACTACGAAACCAACCACTTGTTCACGAGAACGAAATGCACTTACATGGA 180
129 TGTACAACTACGAAACCAACCACTTGTTCACGAGAACGAAATGCACTTACATGGA 188
181 TTGTCAAAACCAAAATCTGCAATACGAATTTAAACCAACATTCATGTTCCAAATTG 240
189 TTGTCAAAACCAAAATCTGCAATACGAATTTAAACCAACATTCATGTTCCAAATTG 248
241 GGGTAATGCTTGTGGGTTGGGGTGAACAACGGCTCAACCTCTACCGGTGTGTTATTG 300
249 GGGTAATGCTTGTGGGTTGGGGTGAACAACGGCTCAACCTCTACCGGTGTGTTATTG 308
301 CTAAACCGAGAGGCGCATTTATGGGCTTACAAAGACAAATTTGAACAAGCAATTACTT 360
309 CTAAACCGAGAGGCGCATTTATGGGCTTACAAAGACAAATTTGAACAAGCAATTACTT 368
361 GGTCCCTCAACCAAGCTCAGCATTCGAGTTGGGCTCTTCAAGGAGAGGAAATCTAG 420
369 GGTCCCTCAACCAAGCTCAGCATTCGAGTTGGGCTCTTCAAGGAGAGGAAATCTAG 428
421 CCCCATTCAGAGCCTGTCTTCATGTGTTAACCTTGACGACATTTGTTGGGGAGTGG 480
429 CCCCATTCAGAGCCTGTCTTCATGTGTTAACCTTGACGACATTTGTTGGGGAGTGG 488
481 AATATGACAACTGAACCTGTGCTGATGCCATGGCCAGGGCAAAAGTTTGAATGAT 540
489 AATATGACAACTGAACCTGTGCTGATGCCATGGCCAGGGCAAAAGTTTGAATGAT 548
541 TGCAGACAGTTGAGGCTTACATGAAATCCATGCTTCACTCCCGGAACTATGACC 600
```

```
549 TCGAGAGCAGTGGAGCGCTTACATGAAATCCATGCTTCCACTCCCCGGAATCTATGACC 608
601 CGGATTTCTATGCTGGCAACCAAGAGAGCTGGCCAAACGTCATGAGGGCAAAAGC 660
609 CGGATTTCTATGCTGGCAACCAAGAGAGCTGGCCAAACGTCATGAGGGCAAAAGC 668
661 AAGAGCAATTCACAAATCATCAAAAGAGATCAAGGCGTTTAAAGGAGCACAAGTGG 720
669 AAGAGCAATTCACAAATCATCAAAAGAGATCAAGGCGTTTAAAGGAGCACAAGTGG 728
721 ACAAGGTGTGTACTGTGAGCTGCCACACAGAGAGTACAGTAATTTGTGTGGGCC 780
729 ACAAGGTGTGTACTGTGAGCTGCCACACAGAGAGTACAGTAATTTGTGTGGGCC 788
781 TTAATGACACCAATGAGAAATCTCTTGGCTGTGTGACAGAAATGAGGCTGAGATTTCTC 840
789 TTAATGACACCAATGAGAAATCTCTTGGCTGTGTGACAGAAATGAGGCTGAGATTTCTC 848
841 CTTCACCTTGTATGCGCATTTGCTGTGTATGAGAAAATGTTGCTTTCATTAATGAGGCC 900
849 CTTCACCTTGTATGCGCATTTGCTGTGTATGAGAAAATGTTGCTTTCATTAATGAGGCC 908
901 CTCAGAACACTTTTGTGACAGAGGCTGATGATCTTGCCATCGGAGAGAACACTTTGATTG 960
909 CTCAGAACACTTTTGTGACAGAGGCTGATGATCTTGCCATCGGAGAGAACACTTTGATTG 968
961 GTGAGATGACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTTGTGATTTCCTTG 1020
969 GTGAGATGACTTCAAGAGTGTGACAGCCAAATGAAATCTGTGTTGTGATTTCCTTG 1028
1021 TGGGGGCTGTATTCAGAACCAATCTATGTAGTATGACCAATCTGAGGAAACAAAGATG 1080
1029 TGGGGGCTGTATTCAGAACCAATCTATGTAGTATGACCAATCTGAGGAAACAAAGATG 1088
1081 GTATGATCTTTCGGCTCCACAACTTTCGTTCCAGAGAAATCTCCAGAGCAACGTTG 1140
1089 GTATGATCTTTCGGCTCCACAACTTTCGTTCCAGAGAAATCTCCAGAGCAACGTTG 1148
1141 TTGATGATATGATGCAACAGCAATGCAATCTCTATGAGCTGTGTGAACATTCAGACATG 1200
1149 TTGATGATATGATGCAACAGCAATGCAATCTCTATGAGCTGTGTGAACATTCAGACATG 1208
1201 TTGTTGTTATTAAGTATGTCCTTACGTAAGGAGACAGCAAGAGACCATGATGATGATA 1260
1209 TTGTTGTTATTAAGTATGTCCTTACGTAAGGAGACAGCAAGAGACCATGATGATGATA 1268
1261 CTTCAAGATATTCATGAGGTGGAAGAGACCAATGTTTGGACAAACACATGCGAGATT 1320
1269 CTTCAAGATATTCATGAGGTGGAAGAGACCAATGTTTGGACAAACACATGCGAGATT 1328
1321 CCCTCTAGCTGCTCTATTTATCTTGAATTGATCTTCTTGTGCTGAGCTCAGACATGAA 1380
1329 CCCTCTAGCTGCTCTATTTATCTTGAATTGATCTTCTTGTGCTGAGCTCAGACATGAA 1388
1381 TCGAGTTTAAAGCTGAAAAATGAGGAAAAATTCATCTCATTCACCCAGTTGCTACATCC 1440
1389 TCGAGTTTAAAGCTGAAAAATGAGGAAAAATTCATCTCATTCACCCAGTTGCTACATAC 1448
1441 TCAGCTACCTCAACCAAGGCTCTGTGTGTCACCGGGATACACAGTGGGAAATGCAATTG 1500
1449 TCAGCTACCTCAACCAAGGCTCTGTGTGTCACCGGGATACACAGTGGGAAATGCAATTG 1508
1501 CAAGCAGCTGCAATGCTGGAACCAATATGAGGGCTTGTGTGATTTGGCCCCAGAGA 1560
1509 CAAGCAGCTGCAATGCTGGAACCAATATGAGGGCTTGTGTGATTTGGCCCCAGAGA 1568
1561 AATAATGATTTCTCGAGTACAAAGTGAACATGGAACCGAAGAAATATATGTTGGGGTAG 1620
1569 AATAATGATTTCTCGAGTACAAAGTGAACATGGAACCGAAGAAATATATGTTGGGGTAG 1628
1621 CTTAGCTGAATGTTTATGTTAATTAATATGTTTGTGTTAATTTTGAAGTATTTGAA 1680
1629 CTTAGCTGAATGTTTATGTTAATTAATATGTTTGTGTTAATTTTGAAGTATTTGAA 1688
```

Oy		1681	TGCATACGCTTCATTAAAGCTTTAGAGCGGGGCAATATTCGTACTACAGAACAATGATG	1740
Db		1689	TGCATACGCTTCATTAAAGCTTTAGAGCGGGCATAATTCGTCTTAACTAGGAACATGAATG	1748
Oy		1741	AATGAGTAATATTTTGCTGT	1760
Db		1749	AATGAGTAATATTTTGCTGT	1768
RESULT 3				
AY038802			1739 bp	linear
LOCUS			Glycine max myo-inositol-1-phosphate synthase mRNA, complete cds.	
DEFINITION				
ACCESSION			AY038802	
VERSION			AY038802.1	
KEYWORDS			GI:14764465	
SOURCE				
ORGANISM			Glycine max (soybean)	
			Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
			Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots;	
			Rosidae; Eustosida I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;	
			Glycine.	
REFERENCE			1 (bases 1 to 1739)	
AUTHORS			Hitz,W.D., Carlson,T.J., Kerr,P. and Sebastian,S.	
TITLE			Biochemical and Molecular Characterization of a Mutation that Confers a Decreased Raffinoseaccharide and Phytic Acid Phenotype on Soybean Seeds	
JOURNAL			Unpublished	
REFERENCE			2 (bases 1 to 1739)	
AUTHORS			Carlson,T.J. and Hitz,W.D.	
TITLE			Direct Submission	
JOURNAL			Submitted (06-JUN-2001) Crop Genetics, Dupont Co., P.O.Box 80402, Wilmington, DE 19880-0402, USA	
FEATURES			Location/Qualifiers	
source			1..1739	
			/organism="Glycine max"	
			/mol_type="mRNA"	
			/cultivar="Wayne"	
			/db_xref="taxon:3847"	
			/dev_stage="developing embryo"	
			11..1543	
			/EC_number="5.5.1.4"	
			/function="cyclizes glucose 6-phosphate to L-myo-inositol-1-phosphate"	
			/note="mlPS-1A"	
			/codon_start=1	
			/product="myo-inositol-1-phosphate synthase"	
			/protein_id="AAK72098.1"	
			/db_xref="GI:14764466"	
			/translation="MEIENFVQECNVYETELISOVNYYETTELHNNNGTYOWIV KPKSKVEKRNINHYPRKICWLVNQMGNGNSTLIGVIANREGISMAPTXKIOOANVF GSLOASAIRGSGRGEETIYAPFKSLPMVNPPDIVEGNDISMNILADAMARKVD IDLOQRPIYMESMLPLFGIYDPFIPLANODEERNANNIKGTKORVOQIIKDIAFEK ATKDDPVTAVLTANTERYSNLVGLNDMETHLLAVIDRNEAEISPSTLYAIACWEVNV PIKNGSPONTVPVPGILDIAIRNTLIGDDFSQTKMSVLVFLVGAGIKPISIVS YNHLHGDMNLISAPOTGRSEIKSNSVDNDSNMILYPGGHPHVHVIKYVPAEVE DKSRAMDEYTSSEIFMGCKSTIVLANCEDSLAAPILIIDIYLALISTIEPEAKNEE GKFHSFHVAFTITLSUTRAPLVPETPVNALSKRAMLENIMACVGLAPENNMLE YK"	
CDS				
ORIGIN				
Query Match			97.6%; Score 1717; DB 8; Length 1739;	
Best Local Similarity			100.0%; Pred. No. 0;	
Matches 1717; Conservative			0; Mismatches 0; Indels 0; Gaps 0;	
Oy		44	GTGAAAAATAATGTTTCATCGAGAAATTTAAAGTTAGTGTCTTAATGGAAGTACACGA	103
Db		1	GTGAAAAATAATGTTTCATCGAGAAATTTAAAGTTAGTGTCTTAATGGAAGTACACGA	60
Oy		104	GACTGAGATTCAGTCCGTGTACAACTACGAAAACCCGCAACTGTTCTTCACGGAACAGGA	163
Db		61	GACTGAGATTCAGTCCGTGTACAACTACGAAAACCCGCAACTGTTCTTCACGGAACAGGA	120

QY	164	TGGCACCTATCAGTGGATTGTCAAAACCCTAATCTGTCAATATACGAATTTTAAACCAACAT	223
DB	121	TGGCACTTATCAGTGGATTGTCAAAACCCTAATCTGTCAATATACGAATTTTAAACCAACAT	180
QY	224	CCATGTTCTTAATATAGGGGTAAATGCTGTGGGTGGGGGTGGAAACAACGGCTCAACCT	283
DB	181	CCATGTTCTTAATATAGGGGTAAATGCTGTGGGTGGGGGTGGAAACAACGGCTCAACCT	240
QY	284	CACCGTGTGTTATTTCTAACCGAGAGGGCATTTCAATGGGCTACAAAGACAAGATTCA	343
DB	241	CACCGTGTGTTATTTCTAACCGAGAGGGCATTTCAATGGGCTACAAAGACAAGATTCA	300
QY	344	ACAAGCCAATTAATCTTGGCTCCTCACCAAGCTCAGCTATCCGAGTTGGGTCTTCCA	403
DB	301	ACAAGCCAATTAATCTTGGCTCCTCACCAAGCTCAGCTATCCGAGTTGGGTCTTCCA	360
QY	404	GGGAGAGAAATCTATGCCCATTTCAAGACCTGCTTCCAAATGGTTAAACCTACACGAT	463
DB	361	GGGAGAGAAATCTATGCCCATTTCAAGACCTGCTTCCAAATGGTTAAACCTACACGAT	420
QY	464	TGTGTTTGGGGGATGGGATATACAGCAATGAACCTGGCTGATGCCATGGCCAGGGCAAA	523
DB	421	TGTGTTTGGGGGATGGGATATACAGCAATGAACCTGGCTGATGCCATGGCCAGGGCAAA	480
QY	524	GGTGTGTTGAATCGATTTGGAGAAGCAGTTGAGGCGCTTACATGAAATCAATGCTTCACCT	583
DB	481	GGTGTGTTGAATCGATTTGGAGAAGCAGTTGAGGCGCTTACATGAAATCAATGCTTCACCT	540
QY	584	CCCCGGAATCTATATACCCCGATTTCAATTCGTGCCAACAGAGAGAGGCTCCAAACGT	643
DB	541	CCCCGGAATCTATATACCCCGATTTCAATTCGTGCCAACAGAGAGAGGCTCCAAACGT	600
QY	644	CATCAAGGGCCAAAGCAAGCAGCAAGTTCAACAATCATCAAGAATCAAGGCGTTTAA	703
DB	601	CATCAAGGGCCAAAGCAAGCAGCAAGTTCAACAATCATCAAGAATCAAGGCGTTTAA	660
QY	704	GGAAGCCACCAAGATGGAACAAGGTGTGTATCTGTGAGCTGCCAACACAGAGAGTACAG	763
DB	661	GGAAGCCACCAAGATGGAACAAGGTGTGTATCTGTGAGCTGCCAACACAGAGAGTACAG	720
QY	764	TAAATTTGGTGTGGGCTTAAATGACACCATGAGAAATCTCTTGGCTGCTGTGACACAA	823
DB	721	TAAATTTGGTGTGGGCTTAAATGACACCATGAGAAATCTCTTGGCTGCTGTGACACAA	780
QY	824	TGAGGCTGAGATTTCTCCTTCACCTTGTAAGCATTGCTGTGTATGGAATAATGTTCC	883
DB	781	TGAGGCTGAGATTTCTCCTTCACCTTGTAAGCATTGCTGTGTATGGAATAATGTTCC	840
QY	884	TTTCATTAAATGGAAGCCTCAGAACCTTTTGTATCAAGGCTGATTTGATCTTGCATCGC	943
DB	841	TTTCATTAAATGGAAGCCTCAGAACCTTTTGTATCAAGGCTGATTTGATCTTGCATCGC	900
QY	944	GAGAAACACTTTGATTGGTGGAGATGACTTCAMAGTGTGCAGCCAAATGAAATCTGT	1003
DB	901	GAGAAACACTTTGATTGGTGGAGATGACTTCAMAGTGTGCAGCCAAATGAAATCTGT	960
QY	1004	GTTGTTGATTCTCTTGGGGGGCTGGATACAGCCAAACCTATATAGCTATTAACAACA	1063
DB	961	GTTGTTGATTCTCTTGGGGGGCTGGATACAGCCAAACCTATATAGCTATTAACAACA	1020
QY	1064	TCTGGGAAACATGATGGTATGAAATCTTTCGGGCTCCAAACTTCCGTTCCAAAGGAAAT	1123
DB	1021	TCTGGGAAACATGATGGTATGAAATCTTTCGGGCTCCAAACTTCCGTTCCAAAGGAAAT	1080
QY	1124	CTCCAAAGCAACGTTGTGATGATATGATCAAGCAATGCCATCTCTATGAGCCTGG	1183
DB	1081	CTCCAAAGCAACGTTGTGATGATATGATCAAGCAATGCCATCTCTATGAGCCTGG	1140
QY	1184	TGAACATTCACACATGTTGTGTTATTAAGTATGTGCTTACCTTAGGGACAGCAAG	1243
DB	1141	TGAACATTCACACATGTTGTGTTATTAAGTATGTGCTTACCTTAGGGACAGCAAG	1200

QY	1244	AGCCATGGAATGATGATCACTTTCAGAGATATTTCAATGGGTGGAAAAAGACCAACATTTGTTTGGCA	1303
Db	1201	AGCCATGGAATGATGATCACTTTCAGAGATATTTCAATGGGTGGAAAAAGACCAACATTTGTTTGGCA	1260
QY	1304	CAACATATGCGGAGATTCCTCTTTCAGTCTGCTCTATATATCTTGGACTTGGTCTCTTGGC	1363
Db	1261	CAACATATGCGGAGATTCCTCTTTCAGTCTGCTCTATATATCTTGGACTTGGTCTCTTGGC	1320
QY	1364	TGAGCTCAGCACTAGATATCGAGTTTAAAGCTGAATAATAGGGAAAAATTCACACTATTTCCA	1423
Db	1331	TGAGCTCAGCACTAGATATCGAGTTTAAAGCTGAATAATAGGGAAAAATTCACACTATTTCCA	1380
QY	1424	CCGAGTTGCTACATCTCTAGCTACCTGACCAAGGCTCTCTGTGTTCCACCGGGTACACC	1483
Db	1381	CCGAGTTGCTACATCTCTAGCTACCTGACCAAGGCTCTCTGTGTTCCACCGGGTACACC	1440
QY	1484	AGTGTGAATGATTTGTCAAAGCAGCGCTGCAATGCTGGAAAAATTAATGAGGGCTTGTGT	1543
Db	1441	AGTGTGAATGATTTGTCAAAGCAGCGCTGCAATGCTGGAAAAATTAATGAGGGCTTGTGT	1500
QY	1544	TGGATTGGCCCCAGAGAAATACATGATTTCTGAGATCAAGGAAGCAATGGGACCGGAGAA	1603
Db	1501	TGGATTGGCCCCAGAGAAATACATGATTTCTGAGATCAAGGAAGCAATGGGACCGGAGAA	1560
QY	1604	TAAATATGTTGGGGTAGCCCTAGCTGAATGTTTATGTATTAATATATATGTTTCTTATATTT	1663
Db	1561	TAAATATGTTGGGGTAGCCCTAGCTGAATGTTTATGTATTAATATATATGTTTCTTATATTT	1620
QY	1664	TTGCAAGTGTATTTGAATGATCATAGCTTCATTAATGCTTTTAGAGCGGGCATATATCTGTT	1723
Db	1621	TTGCAAGTGTATTTGAATGATCATAGCTTCATTAATGCTTTTAGAGCGGGCATATATCTGTT	1680
QY	1724	TACTTGGAAACATGAATGAATGTAGATTAATTTTGTGT	1780
Db	1681	TACTTGGAAACATGAATGAATGTAGATTAATTTTGTGT	1717

RESULT 4			
LOCUS	BD075269	1533 bp	DNA PAT 27-AUG-2002
DEFINITION	Soybean plant producing seeds with reduced levels of raffinose saccharides and phylic acid.		
ACCESSION	BD075269		
VERSION	BD075269.1	GI:32620872	
KEYWORDS	JP 2001519665-A/4.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1533)		
AUTHORS	Hitz,W.D. and Sebastian,S.A.		
TITLE	Soybean plant producing seeds with reduced levels of raffinose saccharides and phylic acid		
JOURNAL	Patent: JP 2001519665-A 4 23-OCT-2001;		
COMMENT	BI DU PONT DE NEMOURS AND CO OS Soybean line LR33 PN JP 2001519665-A/4 PD 23-OCT-2001 PF 07-APR-1998 JP 1998543012 PR 08-APR-1997 US 08/835751 PI WILLIAM DEAN HITZ,SCOTT ANTHONY SEBASTIAN PC C12N15/52,C12N15/82,C12N15/11,C12N9/90,A01H5/00 CC Strandedness: Double; CC Topology: linear; CC Soybean plant producing seeds with reduced levels of raffinose CC saccharides CC and phylic acid FH Key Location/Qualifiers FT CDS 1..1533. Location/Qualifiers 1..1533 /organism="unidentified" /mol_type="genomic DNA"		
FEATURES			
Source			

ORIGIN /db_xref="taxon:32644"

Query Match 87.0%; Score 1531.4; DB 6; Length 1533;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1532; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY	54	ATGTTTCATCGAAGATTTTAAAGTTGAGTGTCTATATGAAGTACACCGAAGCTGAGATT	113
DB	1	ATGTTTCATCGAAGATTTTAAAGTTGAGTGTCTATATGAAGTACACCGAAGCTGAGATT	60
QY	114	CAGTCCGTGTACAACTACGAAACCAACCGAATTGTTCAAGAAACGGAATGGCACTTAT	173
DB	61	CAGTCCGTGTACAACTACGAAACCAACCGAATTGTTCAAGAAACGGAATGGCACTTAT	120
QY	174	CAGTGGATGTGCAAACCCAAATCTGTCAAATACAAATTTAAACCAATCATGTTCTCT	233
DB	121	CAGTGGATGTGCAAACCCAAATCTGTCAAATACAAATTTAAACCAATCATGTTCTCT	180
QY	234	AAATTTAGGGGTAATAGTCTGTGGGTTGGGGGTGAAACCAACGGCTCAACCCCTACCGGTGT	293
DB	181	AAATTTAGGGGTAATAGTCTGTGGGTTGGGGGTGAAACCAACGGCTCAACCCCTACCGGTGT	240
QY	294	GTTATTTGCTAACGAGAGGGCATTTTCATGGGCTTCAAAAGACAAGATTCAACAGCCAT	353
DB	241	GTTATTTGCTAACGAGAGGGCATTTTCATGGGCTTCAAAAGACAAGATTCAACAGCCAT	300
QY	354	TACTTTGGCTCCCTCACCCAGAGCCTCAGCTATCCGAGTTGGGTCTCTTCCAGGAGAGAA	413
DB	301	TACTTTGGCTCCCTCACCCAGAGCCTCAGCTATCCGAGTTGGGTCTCTTCCAGGAGAGAA	360
QY	414	ATCTATGGCCCATTTCAAGAGCCTCTTCCAAATGGTTAAACCTGACGACATTTGTGGG	473
DB	361	ATCTATGGCCCATTTCAAGAGCCTCTTCCAAATGGTTAAACCTGACGACATTTGTGGG	420
QY	474	GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTGTTTAC	533
DB	421	GGATGGGATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGGCAAGGTGTTTAC	480
QY	534	ATCGATTTTGGAGAAGCAGTGAAGGCCCTTAACATGGAATCCATNGCTTCCATCCCGGAATC	593
DB	481	ATCGATTTTGGAGAAGCAGTGAAGGCCCTTAACATGGAATCCATNGCTTCCATCCCGGAATC	540
QY	594	TATGACCCGGATTTTCATTGCTGCTGCCAACGAAGAGAGGTGCCAACAGTCAACAGGAC	653
DB	541	TATGACCCGGATTTTCATTGCTGCTGCCAACGAAGAGAGGTGCCAACAGTCAACAGGAC	600
QY	654	ACAAAGCAAGACAGCAAGTTCACAAATCATCAAGACATCAAGGCGTTTAAAGAACCAACC	713
DB	601	ACAAAGCAAGACAGCAAGTTCACAAATCATCAAGACATCAAGGCGTTTAAAGAACCAACC	660
QY	714	AAAGTGCACAAAGTGTGTTGTACTGTGCACTGCCAACACAGAGAGTACAGTAATTTGGTT	773
DB	661	AAAGTGCACAAAGTGTGTTGTACTGTGCACTGCCAACACAGAGAGTACAGTAATTTGGTT	720
QY	774	GTTGGGCGCTTAATGACCAATGGAANAATCTGTGGCTGTGTGACAAATAATGAGGCTGAG	833
DB	721	GTTGGGCGCTTAATGACCAATGGAANAATCTGTGGCTGTGTGACAAATAATGAGGCTGAG	780
QY	834	ATTTCCTCTTCCACTTGTATATGCAATGCTTGTGTATGAAATGTTCCTTTTCATTAAAT	893
DB	781	ATTTCCTCTTCCACTTGTATATGCAATGCTTGTGTATGAAATGTTCCTTTTCATTAAAT	840
QY	894	GGAGCCCTCAAGAACATCTTTGTATCCAGGGCTGATTGATCTTGCCATTCGGAGAACACT	953
DB	841	GGAGCCCTCAAGAACATCTTTGTATCCAGGGCTGATTGATCTTGCCATTCGGAGAACACT	900
QY	954	TTGATTTGTGGAGATGACTTCAAGAGTGTGACGCAAAATGSAATCTGTGTTGGTTGAT	1013
DB	901	TTGATTTGTGGAGATGACTTCAAGAGTGTGACGCAAAATGSAATCTGTGTTGGTTGAT	960
QY	1014	TTCTTTGTGGGGGTGGTATCAAGCCAAACATCTATAGTCAATTCAACCAATCTGGGAAC	1073

Db 961 TTCCTTGGGGGGCTGATATCAACCAATCTATATGATTCAGTTCACCAATCTGGGAAAC 1020
QY 1074 AATGATGATGATATCTTTCGGCTCCACCAACTTTCGGTCCAGAAATCTCCAAGAC 1133
Db 1021 AATGATGATGATATCTTTCGGCTCCACCAACTTTCGGTCCAGAAATCTCCAAGAC 1080
QY 1134 AACGTTGTTGATGATATGCTCAACAGCAATGCCATCTCTATAGACCTGGTGAACATCA 1193
Db 1081 AACGTTGTTGATGATATGCTCAACAGCAATGCCATCTCTATAGACCTGGTGAACATCA 1140
QY 1134 GACCATGTTGTTGTTATATGATATGCTCTTTCATAGGGGAAACAGAAAGCCCATGAT 1253
Db 1141 GACCATGTTGTTGTTATATGATATGCTCTTTCATAGGGGAAACAGAAATGACCATGAT 1200
QY 1254 GAGTACACTTCAGAGATATTCATGAGTGGTGAAGAGACCATGTTTTCACCAACATGC 1313
Db 1201 GAGTACACTTCAGAGATATTCATGAGTGGTGAAGAGACCATGTTTTCACCAACATGC 1260
QY 1314 GAGGATTCCTCTTGAAGCTGCTCTATATATCTTGAAGCTTCTTCTGAGCTCAGC 1373
Db 1261 GAGGATTCCTCTTGAAGCTGCTCTATATATCTTGAAGCTTCTTCTGAGCTCAGC 1320
QY 1374 ACTAGATGAGTATTAAGCTGAAATGAGGAAATTCATCTATCCAGCCAGTTGCT 1433
Db 1321 ACTAGATGAGTATTAAGCTGAAATGAGGAAATTCATCTATCCAGCCAGTTGCT 1380
QY 1434 ACCATCTCAGCTACCTCCACCAAGGCTCTCTGTTCCACCGGATACACAGTGGTGAAT 1493
Db 1381 ACCATCTCAGCTACCTCCACCAAGGCTCTCTGTTCCACCGGATACACAGTGGTGAAT 1440
QY 1494 GCATTGTCAAAGCAGGTCGAAATGCTGAAAAATATATGAGGCTTGTGTGATTTGGCC 1553
Db 1441 GCATTGTCAAAGCAGGTCGAAATGCTGAAAAATATATGAGGCTTGTGTGATTTGGCC 1500
QY 1554 CCAGAGATTAATATGATTTCTCGATTCAGATGA 1586
Db 1501 CCAGAGATTAATATGATTTCTCGATTCAGATGA 1533

RESULT 5
AB032073 1931 bp mRNA linear PIN 08-SEP-1999
LOCUS Nicotiana paniculata NpINPS1 mRNA for myo-inositol-1-phosphate
DEFINITION Nicotiana paniculata NpINPS1 mRNA for myo-inositol-1-phosphate
synthase, complete cds.
ACCESSION AB032073
VERSION AB032073.1 GI:5834499
KEYWORDS myo-inositol-1-phosphate synthase.
SOURCE Nicotiana paniculata
ORGANISM Nicotiana paniculata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 (bases 1 to 1931)
AUTHORS Hashimoto, A., Yamada, S. and Komori, T.
TITLE Myo-inositol-1-phosphate synthase
JOURNAL Published Only in DataBase (1999)
REFERENCE 2 (bases 1 to 1931)
AUTHORS Hashimoto, A., Yamada, S. and Komori, T.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1999) Akiko Hashimoto, Japan Tobacco Inc., Plant
Breeding and Genetics Research Laboratory, 700 Higashibara, Iwata,
Toyoda-cho, Shizuoka 438-0802, Japan
(E-mail: Akiko.Hashimoto@pbgrl.jti.co.jp, Tel: +81-538-32-7116,
Fax: +81-538-33-6046)
FEATURES
source location/Qualifiers
1..1931
/organism="Nicotiana paniculata"
/mol_type="mRNA"
/db_xref="taxon:62141"
1..1931
/gene="NpINPS1"
92..1624
/gene="NpINPS1"

ORIGIN
Query Match 63.7%; Score 1120.4; DB 8; Length 1931;
Best Local Similarity 83.0%; Pred. No. 4.2e-291;
Matches 1277; Conservative 0; Mismatches 261; Indels 0; Gaps 0;
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="BAAG4084.1"
/db_xref="gi:5834500"
/translation="MFLENFVESPNVKTSESHSVYDYOTTELVDHEKNGYQWTV
KPKVKYEFKTDVHPKLGWLVGMGNGNSTLLGCTIARREGISMAITKDVQAOANVP
GSLTQASTIRVGSFNGEITAFPRKSLIPKVPDVTGGMISPMNLDAKARAVFD
IDLQKOLRPVMSVPLPGIYDPFIANQSSRANVYIKTKKEQIDQIKDIEFKE
KNKVDKVVLTANTERYSNVNVGLNDTMEMLFASVDNRNEAISPSTLYACILENV
PFINSGPONTFVPLIDIAIKNTLIGDDPFGSQTMMKSVLVDLFGAGIKPTISVS
YNHIGNDGMNLSAPOPFRSKEISKSVNVDMSNMLIVERGHPHVVYIKVPPV
GDSRKANDERTYSEIFMGSKRTIYVANTCEBSLLAARPIIDLVTLAEISTRTOLKAE
GKFSFHPVATILSYLTPALVPPTPVNALSIRKAMLENITLACVGLADENNMLE
YK"
51 ATAAATGTCATCGAGAATTTTAAAGTTGAGTGTCTTAATGTGAATACCCGAGACTGAG 110
89 AAAATGTTATTTAGAAATTTTAAAGTTGAGAGCCCAACGTTAATGATACCGAAAGTGA 148
111 ATTCAGTCCGTGTACACTACGAAACCAACGAACTTGTTCACGAGAACAGAAATGCAAC 170
149 ATTCACTGTGTATGATATTAATCAACCACTGATGATTAATGATGAGAAAAATGAGACA 208
171 TATCAGTGGATTTGTCAAACCCAAATCTGTCAAAATGCAATTTAAACCAATCCATGTT 230
209 TATCAATGAGACCGTCACACCTTAAGCTGTCAAAATATGATGTCAGACTGATGTTATGTT 268
231 CTTAAATTAAGGGGTATGCTTGTGGGTTGGGGTGGAAACAACGGCTCAACCTCACCGGT 290
269 CCAAAATTAAGGGGTATGCTTGTGGGTTGGGGTGGAAACAATGTTGATCACTTACCGGT 328
291 GGTGTTATTTGCTAACCGAGGCGATTTTCATGGGTCACAAAGACACAGATTTCAACAGCC 350
329 GGTGTTATTTGCTAACAGAGAAAGAAATTTTCATGGGTCACAAAGATTAAGTTCACAGCC 388
351 AATTACTTGGCTCCCTCACCACCAAGCCCTACGTTATCCGAGTTGGGCTCTTCCAGGAGAG 410
389 AATTACTTGGCTCTCTTACCTCAGGCTTCTACTATTCAGATTTGGGCTCTTTCAGAGAGAA 448
411 GAAATCTATGCCCCCTTCAAGAGCTGCTTCCAAATGTTAAACCTTGACGAGATTTGTTT 470
449 GAGATCTATGCTCATTTTAAAGCTCTCTTCCAAATGTCATTCAGATGAGATGATGTTT 508
471 GGGGATGGGATATTCAGCAACATGAACCTGGCTGATGTCATGGCCAGGCGCAAGGTTT 530
509 GGAGGATGGGACATCAGCAGATGAATTTAGCAGATGTCATGGCCAGGCGCTTAAAGGATTT 568
531 GACATCGATTTGCAAGACAGTTGAGGCTTACATGAAATCCATGCTTCCACTCCCCGGA 590
569 GATATTTGATCTTCAAAACAGTTGAGGCTTACATGAAATCTTATGCTCCACTCCCTGCT 628
591 ATCTATGACCCCGGATTTGATGCTGCCAACCAAGAGAGAGCTGCAACAGTCAATCAAG 650
629 ATCTATGACCCCGGATTTGATGCTGCCAACCAAGGATCACTGCAACAGTCAATCAAG 688
651 GGCACAAAGCAAGCAAGTTCACCAATATCTCAAGACATCAAGACGCTTTAAGGAGCC 710
689 GGAACCAAGAAAGCAATATGATCAATATCTTAAAGATTTAAGGAGTTTAAAGGAAAG 748
711 ACCAAAGTGACAGAGTGTGTACTGTGAGTGCACACAGAGAGTACATGATTTTG 770
749 AACAAAGTGACAGAGTGTGTACTGTGAGTGCACACAGAGAGTACATGATTTTG 808
771 GTTGGGGCTTTAATGACACCATGAGAGATCTCTTGGCTGTGAGACAGAAATGAGCT 830
809 GTTGTGACCTTAATGACACTATGAGAAACCTCTTTGCTTGTGAGACAGAAATGAGCT 868
831 GAGATTTCTCTTCCACCTTGTATGCCATGCTGTGTATGAGAAATGTTCTTTCAAT 890


```

Db      869 GAAATATCTCTCCACTTTGTATGCTATTCCTGCACTTTGAAAATGCGCTTTATT 928
Qy      891 AATGGAAGCCCTCAGAACACTTTTGTACAGAGGCTGATGTATCTTGCCATCGGAGAAC 950
Db      929 AATGGAAGCCCCAGAACACTTTGTGCCAGGCTCATTTATTTGGCCATCAAGAAAGAC 988
Qy      951 ACTTTATTTGTGTAATGACTTCAAGAGTGTGACAGCCAAATGAAATGTGTGTT 1010
Db      989 ACATTATTTGTGTAATGACTTCAAGAGTGTGACAGCCAAATGAAATGTGTGTT 1048
Qy      1011 GATTTCTGTGAGGCTGATCAAGCCAACTATAGTCAGTTACAAACCATCTGGGA 1070
Db      1049 GATTTCTGTGAGTGTGATCAAGCCAACTATAGTCAGTTACAAACCATCTGGGT 1108
Qy      1071 AACATGATGTATGTAATCTTTCGGCTCCAAACTTTCCGTTCCAGGAATCTCCAG 1130
Db      1109 AACATGATGTATGTAATCTTTCGGCTCCAAACTTTCCGTTCCAGGAATCTCCAG 1168
Qy      1131 AGCAAGCTTGTGATGATATGTCACAGCAATGCCATCTCTATAGCTGTGAACAT 1190
Db      1169 AGTAATGTTGTGATATGTCACAGCAATGCCATCTCTATAGCTGTGAACAT 1228
Qy      1191 CCAGACCATGTTGTTTATTAAGTATGTCCTTACGAGGAGCAGCAAGAGAGCATG 1250
Db      1229 CCGACCATGTTGTTTATTAAGTATGTCCTTACGAGGAGCAGCAAGAGAGCATG 1288
Qy      1251 GATGATGACATTCAGAGATATTCATGGGTGAGGAAGCACCATTGTTTGCACACAC 1310
Db      1289 GATGATGACATTCAGAGATATTCATGGGTGAGGAAGCACCATTGTTTGCACACAC 1348
Qy      1311 TGGAGAGATTCCTCTTAAGTGTCTCTATTTATCTTGAATCTGCTCTTCTGAGCTC 1370
Db      1349 TGGAGAGATTCCTCTTAAGTGTCTCTATTTATCTTGAATCTGCTCTTCTGAGCTC 1408
Qy      1371 AGCAGTAGATGAGTTTAAAGCTGAAAATGAGGAAAATTCACATCATTCACCCAGTT 1430
Db      1409 AGTACCCGATTCAGCTCAAGAGCTGAAGAGAGGTAAGTTCCATCTCCACCCCGTG 1468
Qy      1431 GCTACCATCTCAGCTACCTCAGCAAGGCTCTCTGTGTCACCGGGTACACCAATGTG 1490
Db      1469 GCTACCATCTCAGCTACCTCAGCAAGGCTCTCTGTGTCACCAAGGTAACACAGTG 1528
Qy      1491 AATGCTTTGTCAAGCAGCGTGTGATGCTGGAACATATAGGGCTGTGTGATG 1550
Db      1529 AATGCTTTGTCAAGCAGCGTGTGATGCTGGAACATATAGGGCTGTGTGATG 1588
Qy      1551 GCCCAGAGATTAATCATGATTCGAGTACCAAGTGAAG 1588
Db      1589 GCACCAAGAACACATGATTCGAAATCAATGAAG 1626

RESULT 6
E27176 1950 bp DNA linear PAT 18-JUN-2001
LOCUS E27176
DEFINITION Novel INPS gene derived from Nicotiana.
ACCESSION E27176.1 GI:13026394
VERSION E27176.1 GI:13026394
KEYWORDS JP 199918789-A/1.
SOURCE Nicotiana paniculata
ORGANISM Nicotiana paniculata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; Lamiales; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 1950)
AUTHORS Shigehiro Y. and Toshiyuki K.
TITLE Novel INPS gene derived from Nicotiana
JOURNAL Patent: JP 199918789-A 1 13-JUL-1999;
JAPAN TOBACCO INC
OS Nicotiana paniculata
PN JP 199918789-A/1
PD 13-JUL-1999
PF 26-DEC-1997 JP 1997359773

```

```

PR      SHIGEHIRO YAMADA, TOSHIYUKI KOMORI
PI      C12N15/09, C12N5/10, (C12N15/09, C12R1:91), C12N15/00, C12N5/00,
PC      (C12N15/00, C12R1:91)
CC      Strandedness: Double;
CC      Topology: Linear;
FH      Key Location/Qualifiers
FT      CDS 92..1703.
          Location/Qualifiers
          source 1..1950
                  /organism="Nicotiana paniculata"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:62141"

ORIGIN
Query Match 63.7%; Score 1120.4; DB 6; Length 1950;
Best Local Similarity 83.0%; Pred. No. 4,2e-291;
Matches 1277; Conservative 0; Mismatches 261; Indels 0; Gaps 0;

Qy      51 ATATATGTTATGAGAAATTTTAAAGTTGATGCTCTTAATGTGAAGTACACCGAGACTGAG 110
Db      89 AAAATGTTTATGAGAAATTTTAAAGTTGAGAGCCCCCAAGTTAAGTACACCGAAAGTGA 148
Qy      111 ATTCAGTCCGTGTAACATACGAAACACCGAATCTGTTACAGAGAACAGAAATGCGACC 170
Db      149 ATTCAGTCTGTCTATGATTTATCAACACCTGAGTTAGTTCAATGATGAGAAAAATGGACA 208
Qy      171 TATCAGTGAATGTCACCAACCAATCTGTCAATACGAATTTAAACCAACATCCATGTT 230
Db      209 TATCAATGACCGGTCAAGCTTAAGACTGTCAATATATGATTTCAAGCTGATGTATGTT 268
Qy      231 CCTAATTTAGGGTAAATGCTTGTGGTGGGTGGAACACAGGCTCAACCTCACCGGT 290
Db      269 CCCAAATTTAGGGGTAAATGCTTGTGGTGGGTGGAACACATGTTCAACCTTGACCGGT 328
Qy      291 GGTGTTATGTCATACCGAGAGGGCAATTTCAATGGGCTACAAAGGACAAAGATTCACAGGC 350
Db      329 GGTGTTATGTCATACCGAGAGGGAATTTCAATGGGCTACAAAGATTAAGGTGCAACAGGC 388
Qy      351 AATTACTTTGGCTCCCTCAACCCAGGCTCAGTATCCAGTTGGGTCTTCCAGGAGAG 410
Db      389 AATTACTTTGGCTCTCTTAATCAAGGCTTCTACTATTTGAGTTGGGTCTTCAATGAGAA 448
Qy      411 GAAATCTATGCCCATTCAGAGCTGCTTCCAATGTTAACTCTGACGACATTTGTTT 470
Db      449 GAGATCTATGCTCAATTTAAAGCCCTTCCAAATGCTCAATTCAGATACCGTAGTGT 508
Qy      471 GGGGGATGGGATATACGAAACATGAACCTGGTGAATGCAATGGCCAGGGCAAAAGGTTT 530
Db      509 GAGGATGGGACATACGCAATGAATTTAGCAGATGCCATGGCCAGGGCTAAAGGTAAT 568
Qy      531 GACATGATTTTCAGAGCAGTTGAGGCTTACATGAAATCCATGCTTCCATCCCGGA 590
Db      569 GATATGATCTACAAAGCAGTTGAGGCTTACATGAAATCCATGCTTCCATCCCGGT 628
Qy      591 ATCTATGACCCGGAATTTCAATGCTGCTGCAACAAAGAGAGCGTCCAAACAGTCAAG 650
Db      629 ATCTATGACCCGGAATTTCAATGCTGCTGCAACAAAGAGAGCGTCCAAACAGTCAAA 688
Qy      651 GGCACAAAGCAGACCAAGTTCAACAAATCAACAAAGACATCAAGGCGTTTAAAGAGGC 710
Db      689 GGAACCAAGAGAAACCAATATATCAATCAATTAAGATATTAAGGAGGAAAG 748
Qy      711 ACCAAAGTGACAGAGGTTGTACTGTGACCTGCAACACAGAGAGGTACAGTAATTTG 770
Db      749 AACAAAGTGACAGAGGTTGTACTGTGACCTGCAACACAGAGAGGTACAGTAATTTG 808
Qy      771 GTTGTGGGCTTAAATGACACATGAGAAATCTTGGCTGCTGTGACAGAAATGAGGT 830
Db      809 GTTGTGTGACTTAAATGACACATGAGAAACCTTGTGCTGTGTGACAGAAATGAGGT 868
Qy      831 GAGATTTCTCTTCAACCTGTATGCAATGCTGTGTTATGGAATGTTCTTCAAT 890

```

Db	869	GAATATCTCTCCATCCATTGGTATAGCATATGCGTCGANTCTTGAAATATGTCCTTTTATT	928
QY	891	AATGGAAGCCCTCGAACAACCTTTTGACAGGGCTGATTGATCTTGGCATCGAGAAC	950
Db	929	AATGGAAGCCCCCGAACAACCTTTGTGCCAAGGCTCATTTGATTTGGCCATCAAGAAC	988
QY	951	ACTTGAATGGGTGAGATGACTTCAAGATGGGTGAGACCAATGAATCTGTGTGGTT	1010
Db	989	ACATTGATGGTGTGATGACTTTTAAAGTGTCAACCAAAATGAAGTCAGTCTGGTT	1048
QY	1011	GATTTCTTTGGGGGCTGTATCAGCGCAACATCTATAGTCAGTTACCAACATCTGGGA	1070
Db	1049	GATTTCTTTGTGGAGCTGTGATTAAGCAACATCAATGTGAGGTCAACCATTTGGGT	1108
QY	1071	AACAATGATGATGAAATCTTTCCGCTCCACAAACCTTCCGTTCCAGAAATCTCAAG	1130
Db	1109	AACAATGATGAAATGAAATCTGTCTGCCCCCTCAAACTTTCCGGTCAAAGAAATCTCGAAA	1168
QY	1131	AGCAACGTTGTGATGATATAGGTCAACAGCAATGCCATCCCTATGAGCCGTGGAACAT	1190
Db	1169	AGTATGTTGTGATGACATGATGTTTCAAGCAATGCAATCTTTATGAGCCTGGAGAGATC	1228
QY	1191	CCAGACCATGTTGTTGTTATTAAGATATGTGCTTACGTAGGGGACAGCAAGAGACCCTG	1250
Db	1229	CTGACCATGTTGTGTGTGATTAAGTATGTGCCATATGTGGAGACAGCAAGGGCAATG	1288
QY	1251	GATGAGTACACTTCAGAGATATTCATGGGTGAAAGAGACACATTTGTTTTCACAACA	1310
Db	1289	GATGAGTACACATCTGAGATTTTTCATGGGGGAAAGAACACATCTGTTTTCACAATATCT	1348
QY	1311	TGCGAGGATTCCTCTTATAGCTGCTCTATTAATCTTGACCTTGGTCTTCTTGCTGAGCTC	1370
Db	1349	TGTAGGATTCACTTTATAGCTGCTCAATTAATATTTGATTTGGTCTTCTTGCTGAACTTC	1408
QY	1371	AGCACTGAATTCGAGTTTAAGCTGAAATAGAGGAAATTCACATCTACATCCACCTT	1430
Db	1409	AGTACCGCATTCACGCTCAAGAGCTGAAGAGAGGGTAAATGTTCACTCTTCCACCCGCTG	1468
QY	1431	GCTACCATCTCCAGCTACCTCACAAGAGCTCCTCTGATTTCCACCGGGATACCAAGTGTG	1490
Db	1469	GCTACTATCTCCAGCTACCTTACCAAGGCTCCTCTGTATCAACACAGTATACACAGTGTG	1528
QY	1491	AATCATTGTCAAAAGACGCTGCAATGCTGAAAAATATATATGAGGCTTGTGTGGATTG	1556
Db	1529	AATCACACTCTCAAAAGCAGAGGCAATGCTTGAGAAATATATGAGGGCTGTGTGTGACCTT	1588
QY	1551	GCCCCAGAGATTAACATATCTCGAGATACAAAGTGAAG	1588
Db	1589	GCACCAGAGAACATGATTTCTGGAAATTAACAAATGAAG	1626

RESULT 7	
AB009881	
LOCUS	AB009881
DEFINITION	1959 bp mRNA linear PLN 13-JUL-2000
ACCESSION	AB009881
VERSION	AB009881
KEYWORDS	complete cds.
SOURCE	AB009881.1 GI:609265
ORGANISM	myo-inositol 1-phosphate synthase. Nicotiana tabacum (common tobacco)
	Nicotiana tabacum Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicots/ladons; core eudicots; asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE	1 (bases)
AUTHORS	Hara,K., Yagi,M., Koizumi,N., Kusano,T. and Sano,H.
TITLE	Screening of wound-responsive genes identifies an immediate-early expressed gene encoding a highly charged protein in mechanically wounded tobacco plants
JOURNAL	Plant Cell Physiol. 41 (6), 684-691 (2000)
MEDLINE	20399434
PUBMED	10945337
REFERENCE	2 (bases 1 to 1959)

AUTHORS Hara, K., Yagi, M., Koizumi, N., Kusano, T. and Sano, H.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-1997) Kojiro Hara, Nara Institute of Science and Technology

TECHNOLOGY, Plant Molecular Breeding, Yakaku 8316-5, Ikomachi, Nara 630-0101, Japan (E-mail: k-hara@bs.aisc-nara.ac.jp).
Tel: +81-743-72-5653 (ex. 5653), Fax: +81-743-72-5655

source	location/Qualifiers
1. 1959	
	/armenian-III'settian tabrown"

```

/organism="Nicotiana glauca"
/mol_type="mRNA"
/db_xref="taxon:4097"
/tissue_type="leaf"
94..1626

```

```

/codon_start=1
/product="myo-inositol 1-phosphatase"
/protein_id="BA95788.1"
/db_xref="GI:8096266"

```

```

/translation="MRLENNKVSPPVKKITLSSHSYDQITLHLEHKNNGIYQVU/
KPTVYXEFKTDVHVEKLGMLVGMGNGNSTLTGVIAAREGISNATKADQOQANF/
GSLTOSTATRVSGFNGEVIAPFKSLIPWNPDVFGVDIMTADMAAKAVD/
IDLQKLRPMESMVDPLPGIYDPTLAANGSANNVIKTKKKEQIDQIIKIDREFKE

```

KXNVJDKVVLWTINTERYSVVVGLNDIMNENI.PASVDREAEISPTLYIAIACLENNV
PFINSXPONTFVPGLLDLAIKRNLTIGDDFKSGQTKMSLVDFVDEHGAICIKPISIVS
YHLSGNDDGNLSAPQTFRPSKEIKSGNEPHDVSNSNALVPELPHVYVIVKIPVYV
GSKRAMDEYTSIEIFMGQKNITVLANHCESLSLAAPILDLVLIAELSTRIQLKAEGBE

GKHSHHPVATILSYLTAKPLVPGIPVNAALSKQAMLENIIRACGLAPENNMIIE
 YK"
 1959
 polyA_site
 /note="15 a nucleotide"

	ORIGIN	Accession No.	Accession No.	Accession No.
Query Match	63.4%	Score	1115.6;	DB 8; Length 1959;
Best Local Similarity	80.4%	Prod No	A4e-290.	

seq. model similarity	0.75	1.00	NO. DITS	250	
Matches 1307;	Conservative	0;	Mismatches	319;	Indels
					Gaps
					0;

26 CATTCACTTCTTAATCTTTGGAAAAATATGTCATGAGAAATTTAAGCTTGAGTGTCC 85

Db 66 CTTGATCCTTTCTAAAGTTCCCAAAAGTTTATGACACTTTAAGTTTGAGACCC 125

Oy 86 TAAATGTAAGTACACGAGACTGAGATTGAGTCCTGTACACTACGAAACCAAGCACT 145

Db 126 CAAAGTTAAGTACCCGAAAGTAGATTCTCTGTATGATTATCAACCACTGAGTT 185

Oy 146 TGTTCACGAGAACGGAATGGCACCTATCGATGGATTGTCAAAACCCAAATCTGTCAATA 205

Db 186 AGTTCATGAGGAAAAATGGGACTTACCAATGGACTGTCAAGCTTAAGACTGTCAATA 245

DY 206 CGAATTTAAAAACAACATCCATGTTCTTAAATTAGGGGTATGCTTGTGGGTTGGGGTGG 265

Db 246 TGAGTTCAGACTGATGTTTCATGTTCCCAATTAGGGGTTATGCTTGTTGGATGGGGTGG 305

Y 266 AAACAAACGCTCAACCCCTACCGGTGTTGTTATGTTACTAACCGAGAGGGCAATTTTCATGGC 325

326 TACAAAGGACAAGATTTCAACAAGGCAATTACTTTGGTCCCTCAACCCGAAGCTCAGTAT 385
306 AACCAATGTTCAACCTTGACCGGTGGTGTATTGCTAACAGAGAAAGAAATTCATGGGC 365

366 CACCAAGAATGAAGTGCACACAGCAATTACTTTGGCTCTTACTCAGGCTTTACTAT 425

Dd

CCGAGTTGGGTCTTTCATATGAGACATCTATGCTCCATTAAAAAGCCTCCTCCAAT

485

446 GGTAAACCCCTGACGACATTGTTGTTGGGGATGGGATATACGACATCAATGAACCTGGGAGA 505
 486 GGTCAATCCAGATGACGTAGTGTGGAGGATGGGACATCAGCGGCATGAAATTTGGGAGA 545

Dy 506 TGCAATGGCCAAAGGCTGTTCATCGATTGCCAAGAAGTTGAGGCCTTACAT 565
| | | | |
Db 546 TGGCATGGCAAAGGCTTAAGCTATTGATAATTGATCTCACAAGAAGTTGAGGCCTTACAT 605

Q7 566 GGATTCATCTTCCAATCCCCGGAATCATGACCCGATTTCAATGCTGCACCAAGA 625
||||| ||| | ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Dh	606	GGAAATCTATGTCGCCCACTACCTGGGATCTATGATCTCTGATTTTCATTGTCGCTCAACCAAG	665
Qy	626	GGAGCGTCCCAACCAACGTCATCTCAAGGGCACAAAGCAAGCAAGTTCAACAACATCATCA	685
Dh	666	GTCACGTGCCAACCAACGTGATCTAAAGGAACCAAGAAAGAACAAATTTGATCAAAATCATTTA	725
Qy	686	AGACATCAAGGCGTTTAAGGAAGCCACCAAGTGAACAAGTGTGTGTAAGTGTGAATGCG	745
Dh	726	GGATATTTGGGAGTTTAAAGGAAGAACAAAGTGAACAAGGTGTAGTCTGTGTGATGTC	785
Qy	746	CAACACAGAGAGGTACAGTAAATTTGGTTGGGCGCTTAATGACACATAGGAATCTCTT	805
Dh	786	TAACTCTAAAGATACAGTAATGTGTGTGTGGAATTATGATTAAGTAAGAAATCTCTT	845
Qy	806	GGCTGCTGTGAACAAGAAATGAGGCTGAGATTTCTCTTCACTTGTATGCACTGCTTGG	865
Dh	846	TGCTTCTGTGACAGAAATGAAAGCTGAAATATCTCTTCACTTGTATGATTAATGCTGTG	905
Qy	866	TGTTATGGAATAATGTTCTCTTTCATTAATGGAAGCCCTCAAAACATTTTGTACAGGCT	925
Dh	906	CATCTCTGAATAATGTGCGCTTTTATTAATGGAAGCCCAAGAACACCTTTGTGCGAGGCT	965
Qy	926	GATTGATCTTGCAATCGGAGGAACATTTGATTTGTGTGAGATGATCTTCAAGAGTGTCA	985
Dh	966	CATTGATTTGGCATCAAGAGGAACACACTGATGTGTGTGTGATCTTTAAGAGTGTCA	1025
Qy	986	GACCAAAATGAATCTGTGTGTGTGATTTCTTGTGTGGGCGCTGATCAAGCCCAACATC	1045
Dh	1026	AACCAAAATGAAGTACGTGTGTGTGATTTCTTGTGTGAGCTGTATTAAGCCAAATC	1085
Qy	1046	TATAGTCATTTACCAACATCTGGGAAACAATGATGTATGAAATCTTTGGCTCCAAAC	1105
Dh	1086	AATGTGAGCTACCAACCAATTTGGGTAAACAATGATGGAATGAAATCTGTCTGCGCCCTCAAC	1145
Qy	1106	TTTCCGTTCCAAAGAAATCTCCAAAGAACAGTCTTGAATATAGTCAACAGAAATGC	1165
Dh	1146	TTTCCGCTCAAGGAGATCTGAAAGTAATGTGTATGATGACATGATTTCAAGCAAGAACG	1205
Qy	1166	CATCCTCATAGAGCTGTGTGAACATCCAGACCATTTGTGTATTAAGTATGTCCTTA	1225
Dh	1206	CATCTTTTATGAGCTGTGAGAGCACCTGACCATGTGTGTGTGATTAAGTATGTGCCAT	1265
Qy	1226	CGTAAAGGAGACAGCAAGAGAGCCATGATGAGTACATTTCAAGAGATTTTCATGGGTGAAA	1285
Dh	1266	TGTGGGAGACAGCAAAAGGCAATGATGAGTACATCTGAGATTTTCATGGGAGAAA	1325
Qy	1286	GAGCAACCATTTGTTTGACAACAATGCGAGGATTCCTCTTAACTGCTCCTATATATTT	1345
Dh	1326	GAAACCACTTTTGTGCAAAATCTTGTGAAGATTTCTTTTGTGCTCTCAATATATCTT	1385
Qy	1346	GGACTGGTCTCTCTGTGTGAGCTCAGACCTAAGATGAGTTAAAGCTGAAGAAATGAGGG	1405
Dh	1386	GGATTTTGTCTCTTGTGTGAACTCAGTACCGCATTCAGCTCAAAAGCTGAAGGAGAGGG	1445
Qy	1406	AAAAATTCACATCATTTCCACCCAGTTGTCTACCATCTCAGCTAACCTCAACAGAGCTCTCT	1465
Dh	1446	TAAGTTCACATCTTTCCACCCCGGTGGCTAATATCTCTAGCTAACCTTACAAAGGCTCTCT	1505
Qy	1466	GGTTCACCGGGTACACCATGTGTGAATGCAATTTGTCAAAAGACGCTGCAATGCTGAAAA	1525
Dh	1506	GGTACACCAAGATACACCATGTGTGAATGCACTTCAAAAGACGAGGGCAATGCTTGAGAA	1565
Qy	1526	CATATAGAGGCTGTGTGTGATTTGGGCCCAAGAAATTAACATGATTCCTCGAGTCAAGTG	1585
Dh	1566	CATATTGAGGCTGTGTGTGATTTGCAACCAAGAGACAACTGATTTGGAATACAAATG	1625
Qy	1586	AAGCATGGGACCGAAGATTAATATAGTTGGGTAGCTTACGTGATGTTTATGTTATA	1645
Dh	1626	AAGATCGGAAAGCACTCAAGAAAGCTGATCGGCAAGAAACATATATGTTCTTACTCAAGA	1685
Qy	1646	ATATGT 1651	
Dh	1686	AGCTGT 1691	

LOCUS	BT013759	1954 bp	mRNA	linear	PLN 11-MAY-2004							
DEFINITION	Lycopersicon esculentum clone 132642F, mRNA sequence.											
ACCESSION	BT013759											
VERSION	BT013759.1	GI:47105174										
KEYWORDS	FLI CDNA.											
SOURCE	Lycopersicon esculentum (tomato)											
ORGANISM	Lycopersicon esculentum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 1954)											
REFERENCE	Kirkness, E.F., Wang, W. and Vazellie, A. Direct Submission Submitted (11-MAY-2004) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA											
AUTHORS	Location/Qualifiers											
TITLE	1..1954											
JOURNAL	/organism="Lycopersicon esculentum" /mol_type="mRNA" /db_xref="taxon:4081" /clone="132642F" /issue_type="fruit" /note="TZBXY41"											
FEATURES	source											
ORIGIN	62.1%; Score 1092.2; DB 8; Length 1954; Beet Local Similarity 81.5%; Pred. No. 1.98-263; Matches 1265; Conservative 0; Mismatches 288; Indels 0; Gaps 0;											
Query Match												
35	TTAATCTTTGGAAAAAATAATGTTTCATCGAGAAATTTAAGGTGATGTCCTAATGTGAA	94										
88	TTTGTGTTTGGATGATAAAATGTTTATTTGAACTTTAAGGTGAGGCCCAATGTTAA	147										
95	GTACACGAGCTGAGATTGAGTCCGATGTCACATACGAAACACGCGAATCTGTTACGA	154										
148	GTACACGAGAGTAGATTCACTCTGTGTATGATTAACAAACAATGATGTTGTTACATGA	207										
155	GACAGAAATGGCACTATCAGTGGATTTGTCAAACTGATCTGCAATATGCAATTTAA	214										
208	GGAATAAATGGGACATTTCAATGAGACTGTAAAGCCAAAATCTCAAAATATGAATTTCA	267										
215	AACCAATCATCTTCTCTAATTTAGGGGTATGCTTGTGGTTGGGGTGAACAACCG	274										
268	AACGTATGTTCAATGTTCCAAATATGAGATTAATGCTTGTGGATGGGTGAACAATGG	327										
275	CTGAACCTCAACCGATGATTTATGTTAACCAGAGAGGCAATTCATGGGCTACAAAGA	334										
328	TTTACGCTTTCGAGAGGTGTATTTGCTAATATGAGAAATTTCAATGAGGCTACAAAGA	387										
335	CAAGATTCACAAGCCATTAATTTGGCTCCCTCACCCCAAGCCTCACTATCCGATTTGG	394										
388	TAAAGTGAAGAGCAATTAATCTTGTGTTCTTTACACAGGCTCTACTATTTGAGTTGG	447										
395	GTCCTTCACGAGAGAGAAATCTATGCCCCATTCAGAGCCTGCTTCAATGTTAAACC	454										
448	GTCCTTTCAATGGAGAGAGATCTATGCTCCCTTTAAGACATCTTCCCATGTGCAATCC	507										
455	TGAAGCATTTGTTTGGGGAGATGGGATATCAGCAATGAACTGTGCTGATGCAATGGC	514										
508	AGATGATGATTTTGGAGAGATGGGACATCGACATGAAATTTGGCAATGCAATGGC	567										
515	CAGGCAAAAGGTGTTTGACATGCAATTTGCAAGACAGTGAAGCCTTACATGCAATTCAT	574										
568	AAGAGCTAAAGTTTGTGACATTTGCTTCAAAATGCAACTGAGGCGGTGATGCAATTCAT	627										
575	GCTTCCACTCCCGGAAATCTATGACCCGGAATTTTCATGCTGCTCCAACTAAGAGAGCTGC	634										
628	GGTCCACTGCTGTGATCTATGACCTCTGATTTTCATTTGCTGCAAACTCAAGCTTCACTGC	687										

635 CAACACGTCATCAAGGGCACAAGACAGCAAGTTCACAAATCATCAAAAGACATCA 694
Db TAACATGTCATCAAGGAACCAAGAAACAAATATGATCAATATTAAGATATTAAG 747
Qy GGGCTTTAAGGAAGCCACCAAGTGAACAAGTGTGTACTGTGTGACCTGCCAACACGA 754
Db GGAGTTTAAAGGAACCAAGTGAACAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 807
Qy GAGGTATGATTAATTTGGTGTGGGCTTATGACACCATGAGAAATCTCTGTGGTGTGT 814
Db AAGATATGACAGCATGT 867
Qy GGACGAAATGAGGCTGAGATTTCTCTTCCACTTTGATGTCATGTCATGTCATGTCATG 874
Db AATGTATCCCTTCATCAATGAGAACCCCAAAATCTTTGTGTGTGTGTGTGTGTGTGT 987
Qy TGGCATCGGAGAAACACTTTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 994
Db GGCATCAAGAGAAACCTTTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1047
Qy GAAATCTGT 1054
Db GAAGTCAGT 1107
Qy TTACAAACCTCTGT 1114
Db CTACAAACCTCTGT 1167
Qy CAAGAAATCTCAAGAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1174
Db AAGAGAGATCTGAAAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1227
Qy TGAGCTGT 1234
Db TGAGCTGT 1287
Qy CAGCAAGAGAGCCATGT 1294
Db TGACAAAGAGGCAATGT 1347
Qy TGT 1354
Db TGTATTTGCAAACTCTGT 1407
Qy CTTTCTGT 1414
Db GCTCTGT 1467
Qy CTGATTTCCACCAAGT 1474
Db CTCCTTCAATCCCGTGACCAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1527
Qy GGGTACACCAAGT 1534
Db AGGTACACCAAGT 1587
Qy GGGT 1587
Db GGCTGT 1640

RESULT 9
BT013505 1959 bp mRNA linear PLN 11-MAY-2004
LOCUS Lycopersicon esculentum clone 132193F, mRNA sequence.
DEFINITION BT013505
ACCESSION BT013505
VERSION BT013505.1 GI:47104920
KEYWORDS FLI_CDNA.
SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 1959)
AUTHORS Kirkness, E.F., Wang, W. and Vazeille, A.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
source location/Qualifiers
1..1959
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="132193F"
/tissue_type="fruit"
/note="TZKBF51"

ORIGIN
Query Match 61.8%; Score 1087; DB 8; Length 1959;
Best Local Similarity 80.6%; Pred. No. 4.4e-282;
Matches 1270; Conservative 0; Mismatches 305; Indels 0; Gaps 0;
Qy 36 TAACTTTGTGAAAATAATGTTCTGACGAATTTTAAGTGTGATGTCCTTAATGTGAAG 95
Db 59 TGAATTTTGTGTCAAAATGTTTAAATTTTAAAGTGAAAGCCCAATGTGAAG 118
Qy TACACGAGACTGAGATTCAGTCCGTGTACAACTAGAAACCAACGCACTTTGTACGAG 155
Db 119 TATATGAGATGAATTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 178
Qy 156 AACAGAAATGGACCTATCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 215
Db 179 GAGAGAAATGAACTTATCAATGATGTGTAAAGCTTAAGCTGAATATGAATTTTAA 238
Qy 216 ACCAATCTCATGTTCTTAATTTAGGGTAATGCTTGTGGGTGTGGGTGTGGAAACAACGC 275
Db 239 ACTGATACCCATGTGCAAAAATGGGGGTGTATGCTTGTGTGAATGGGAGAAACAATGT 298
Qy 276 TCACCTTCACCGGAGTGTATTGTCTAACGAGAGGCACTTCAATGGGCTTACAAAGAGC 335
Db 299 TCAACATTTGATCGAGAGTGTATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 358
Qy 336 AAGATTCAACAGCCAAATTTACTTGTGCTCCCTCACCAAGCCTCAGTATCCGAGTTGGG 395
Db 359 AAAGTGAACAAGCCAAATTTATTTGGGTCTCTTACTCAGGACATCAACCATTCGAGTTGGG 418
Qy 366 TCTTTCACGAGAGAAATCTATGCCCATTTCAAGAGCTGCTTCCAAATGTTAAACCT 455
Db 419 TCTTTCAAATGGGAGAAATCTATGCACTTCAAAACCTTCCCAATGCTCAACCA 478
Qy 456 GACGATTTGTGTGGGGAGTGGATATACGACATGAACCTGTGATGCCATGAGCC 515
Db 479 GACGATGTATTTGTGAGATGAGACATTTAGCAACATGAATTTGGCAGATGCTATGGCT 538
Qy 516 AGGCAAGAGTGTGTGACATGATTTGCAAGACAGTTGAGGCTTACATGGAATTCATG 575
Db 539 AGGCTAAAGTTTGTGAAGTTGATCTGCAAAAGCAGTGAAGCCCTTACATGGAATTCATG 598
Qy 576 CTTCACATCCCGGAATCTATGACCCGGAATTTCAATGCTGTCCAAACCAAGAGAGGTGCC 635
Db 599 GTTCCCTCTTCGTGTATGTATGACCTGATCTTCAATTTGGGCTTACCAAGAAAGCAAGTCC 658
Qy 636 AACAGTTCATCAAGGGCACAAGCAAGCAAGTGTCAACAAATCATCAAGACATCAAG 695
Db 659 AACACGTGATCAAGGAACCAAGAAAGCAAGTGTGAACAAATGTTTAAAGATTTTAAAG 718
Qy 696 GCGTTTAAAGAACCAAGATGTGACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 755
Db 719 GAGTTCAAGGAGAAACAGAGTAGACAAAGATAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 778
Qy 756 AGGTACAGTAATTTGTGTGGGCTTAAATGACACCATGAGAAATCTGTGCTGTGTG 815

Db 779 AGGTACAGTAATGTTGTTGGCTTAAATGACACATGAGAAACCTTTTAGCTGTG 838
 Qy 816 GACAGAAATGAGGCTGATTTCTCCCTCCACCTGATGATCCATGCTGTGATGAA 875
 Db 839 GATAGAAATGAGGCTGATTTCTCCCTCCACCTGATGATCCATGCTGTGATGAA 898
 Qy 876 AATGTCCTTTCTTAATGAGAACCTCCAGAACATTTTGTACAGGCTGATGATCTT 935
 Db 899 AATGTCCTTTCTTAATGAGAACCTCCAGAACATTTTGTACAGGCTGATGATCTT 958
 Qy 936 GGCATGCGAGAGAACCTTTGATGTTGATGATGATGATGATGATGATGATGATG 995
 Db 959 GGCATGCGAGAGAACCTTTGATGTTGATGATGATGATGATGATGATGATGATG 1018
 Qy 996 AATGTCCTTTCTTAATGAGAACCTCCAGAACATTTTGTACAGGCTGATGATCTT 1055
 Db 1019 AATGTCCTTTCTTAATGAGAACCTCCAGAACATTTTGTACAGGCTGATGATCTT 1078
 Qy 1056 TACAAACATCTGGGAAACATGATGATGATGATGATGATGATGATGATGATG 1115
 Db 1079 TACAAACATCTGGGAAACATGATGATGATGATGATGATGATGATGATGATG 1138
 Qy 1116 AAGGAAATCTCCAGAGCAAGCTTTGATGATGATGATGATGATGATGATGATG 1175
 Db 1139 AAGGAAATCTCCAGAGCAAGCTTTGATGATGATGATGATGATGATGATGATG 1198
 Qy 1176 GAGCCTGGAACATGACCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1235
 Db 1199 GATCTGGGAGAGACCTTACATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1258
 Qy 1236 ACCAAGAGAGCAGTATGATGATGATGATGATGATGATGATGATGATGATG 1295
 Db 1259 ACCAAGAGAGCAGTATGATGATGATGATGATGATGATGATGATGATGATG 1318
 Qy 1296 GTTTTGCACACATGACGAGATTCCTCTTACGCTCTTATTAATGTTGATG 1355
 Db 1319 GTTTTGCACACATGACGAGATTCCTCTTACGCTCTTATTAATGTTGATG 1378
 Qy 1356 GTTTTGCACACATGACGAGATTCCTCTTACGCTCTTATTAATGTTGATG 1415
 Db 1379 GTTTTGCACACATGACGAGATTCCTCTTACGCTCTTATTAATGTTGATG 1438
 Qy 1416 TCATTCACACATGACGAGATTCCTCTTACGCTCTTATTAATGTTGATG 1475
 Db 1439 TCATTCACACATGACGAGATTCCTCTTACGCTCTTATTAATGTTGATG 1498
 Qy 1476 GGTACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1535
 Db 1499 GGTACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1558
 Qy 1536 GGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1595
 Db 1559 GGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1618
 Qy 1596 CCGAAGATATATATA 1610
 Db 1619 ATGCAATGAGAAA 1633

RESULT 10
 CPINOIG 1978 bp DNA linear PLN 26-JUN-1995
 LOCUS C.paradisi (Macf) INOI gene.
 DEFINITION Z32632.1 GI:602564
 ACCESSION INOI gene.
 VERSION 232632.1
 KEYWORDS Citrus x paradisi
 SOURCE Citrus x paradisi
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Rosales; eudicotids II; Sapindales; Rutaceae; Citrus.
 REFERENCE 1 (bases 1 to 1978)
 AUTHORS Holland,D.

TITLE Direct Submission
 JOURNAL Submitted (08-APR-1994) Doron Holland, Fruit Tree Breeding and
 Genetics, Volcani Research, Organisation Center, Bet Dagan, 50250,
 Israel
 FEATURES
 source Location/Qualifiers
 1..1978
 /organism="Citrus x paradisi"
 /mol_type="genomic DNA"
 /db_xref="taxon:37656"
 /clone="C38"
 /issue_type="fully expanded leaves"
 /dev_stage="juvenile plants"
 188..1711
 /gene="INOI"
 188..1711
 /gene="INOI"
 CDS
 /gene="INOI"
 /codon_start=1
 /protein_id="CAA83565.1"
 /db_xref="GI:602565"
 /db_xref="GOA:P42802"
 /db_xref="UniProt/Swiss-Prot:P42802"
 /translation="MTENFKVESPVKXTDHEHSHVYDEYETTELVENRNGTYOMIV
 KRTVYBEKIDVHPKLGMLVGMGNNGSTLTGVIANRBIQWATDQYQANVF
 GSLTQSAIRVGSYNGEELIYAPFKSILPMVNPDDIYFGMDISDMNLADAMARVFD
 IDLOLRPYMESMVPLEPIYDPIAANOGSRANVIVIGTKKQWLOIILKIDREKBE
 RKXDRVVVLTANTERYSNVIVGNDPESLASLDKREAEISBETLYAIAQVLENI
 PRINSPOVTFVGLIDLAIRNCLIGDDFKSGGQTKMSVLDFLVGAGIKFTSIVS
 VNHLENDQMNLSAPQTFERSKEISKNVVDVSSVNFPMGLVNRPRAIKVPYVAI
 ERAMDEYSEIFMGKSTVILHNTCEDSLAIPILIDLVLELSRIOLKKEGBEKF
 HSFHIVATILSYLTKRAPLVPPGTVPVNALSKORAMLENILRACVGLAPENNMILBYK"

ORIGIN
 Query Match 61.0%; Score 1073.4; DB 8; Length 1978;
 Best Local Similarity 82.0%; Pred. No. 2.1e-278;
 Matches 1275; Conservative 0; Mismatches 271; Indels 9; Gaps 3;
 Qy 36 TATCTTTTGAAATAATATGTCATCGAATTTTAAAGTTGAGTGTCTAATGGAAG 95
 Db 170 TAAATTTTCTGGGAAATGTTTATCGAAATTTCAAGTTGAGGCCAAATGTGAAG 229
 Qy 96 TACACGAGACTGAGATTCAGTCCGTGTACAATAAGAACCAACCACTTGTTCACGAG 155
 Db 230 TACAGAGATCAGAGATTCAGTCTGTGTACATTAATGAACCACTGAGCTTGTCTATGAG 289
 Qy 156 AACAGAGATGCACTATCAGTGTGATGTCGAACCCAAATCTGTCAATATGCAATTTAA 215
 Db 290 AACAGAGATGCACTATCAGTGTGATGTCGAACCCAAATCTGTCAATATGCAATTTAA 349
 Qy 216 ACCAATCTCAGTTCCTTAATTAAGGGGTAAAGCTTGTGGGTGGGTGGAACCAACGGC 275
 Db 350 ACCGATGTCATGCTCCCTTAACCTTGGGTAACTTGTGGGTGGGTGGAACCAACGGC 409
 Qy 276 TCAACCTTCACCGGTGTGTATGCTTAAACGAGAGGCAATTTCAAGGCTCAAAAGAC 335
 Db 410 TCACTTCACCGGTGTGTATGCTTAAACGAGAGGCAATTTCAAGGCTCAAAAGAC 469
 Qy 410 TCACTTCACCGGTGTGTATGCTTAAACGAGAGGCAATTTCAAGGCTCAAAAGAC 469
 Db 336 AAGATTCAACCAATTAATTAATTAAGGGGTAAAGCTTGTGGGTGGGTGGAACCAACGGC 395
 Db 470 ACTGTGACAGCAAGCAATTAATTAATTAAGGGGTAAAGCTTGTGGGTGGGTGGAACCAACGGC 529
 Qy 396 TCCCTTCAGAGAGAGAAATCTATGCCCATTCAGAGGCTGCTTCAATGTTAACTT 455
 Db 530 TCTTCAATGAGAGAGATTTATCTCCATTCAGAGGATTTCTCCATGTTGAACCA 589
 Qy 456 GACGACATTTGTTGGGGATGAGATATCAGCAACATGAACCTGAGTATGAGGCT 515
 Db 590 GATGACATTTGTTGGGGATGAGATATCAGCAACATGAACCTGAGTATGAGGCT 649
 Qy 516 AGGCGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 575
 Db 650 AGGCGTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 709
 Qy 576 CTTCACATCCCGGAAATCTATGACCCGATTTCAATTTGCTGCTCAACCAAGAGAGGCTGCC 635

```

Db      710 GTCCACCTTCGGAATCATATGACCTGATTTCTATGCTGCCAATCAAGGGCGGTGCC
Qy      636 AACACGCTCATAGGAGCAAAAGCAAGACAGTTCACAAATCATCAAAACACTCAAG
Db      770 AACACGCTCATAGGAGCAAAAGCAAGACAGTTCACAAATCATCAAAATATTTAGA
Qy      696 GCGTTTAAGAAAGCCACCAGTGAACAAGGTGTGTACTGTGAGCTGCCAACAGAG
Db      830 GAGTTTAAGAGAAAAACAGGTGATAGGCTTGTGTACTGTGAGCTGCCAACAGAG
Qy      756 AGGTACAGTAATTTGGTGGGCTTAATGACACATGAGAAATCTCTGGCTGCTG
Db      890 AGGTACAGTAATTTGGTGGGCTTAATGACACATGAGAAATCTCTGGCTGCTG
Qy      816 GACAGAAATGAGGCTGAGATTTCTCTTCACCTGATGACATGCTGTTGTTATGAA
Db      950 GACAAAAAGAAAGCTGAGATTTCTCTTCACCTGATGATGCTGTTGTTATGAA
Qy      876 AATGTTCTTTCATTAATGAAAGCCCTCAGAACCTTTGTACAGGGCTGATGATCTT
Db      1010 AATATCTCTTTCATTAATGAAAGCCCTCAGAACCTTTGTACAGGGCTGATGATCTT
Qy      936 GCCATCGAGAGAGACCTTTGATGATGATGATGATGATGATGATGATGATGATG
Db      1070 GGTATTAGAGAGACCTTTGATGATGATGATGATGATGATGATGATGATGATG
Qy      996 AATCTGCTGTTGTTGATTTCTTGTGGGGCTGATTAACCAACATCTATAGTCA
Db      1130 AATCTGCTGTTGTTGATTTCTTGTGGGGAGAGATTAACCAACATCTATAGTCA
Qy      1056 TACAAACATCTGAGAAACATGATGATGATGATGATGATGATGATGATGATGATG
Db      1190 TACAAACATCTGAGAAACATGATGATGATGATGATGATGATGATGATGATG
Qy      1116 AAGAAATCTCCAGAGCAACGTTGTGATGATGATGATGATGATGATGATGATG
Db      1250 AAGAGATTTCCAAAAGTAATGTTGTGATGATGATGATGATGATGATGATGATG
Qy      1176 GAGCTGCTGGAACATCCAGACCATGTTGTTTATTAATGATGATGATGATGATG
Db      1309 GGGCTGCTGGAACATCCAGACCATGTTGTTTATTAATGATGATGATGATG
Qy      1236 AACAGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG
Db      1361 ATCGAGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG
Qy      1296 GTTTTGCACACATGCGAGATTCCTCTTAGCTGCTCTATTAATCTTGAACCTTGT
Db      1421 GGTCTGCACACATGCGAGATTCCTCTTAGCTGCTCTATTAATCTTGAACCTTGT
Qy      1356 GTTCTGCTGAGCTACGACTGATGATGATGATGATGATGATGATGATGATGATG
Db      1481 CTGCTGCTGAGCTACGACTGATGATGATGATGATGATGATGATGATGATGATG
Qy      1416 TCATTTCACCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG
Db      1541 TCTTTTCACCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG
Qy      1476 GGTACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG
Db      1601 GGTACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG
Qy      1536 GCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
Db      1661 GCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG

```

RESULT 11
 AF284065 1845 bp mRNA linear PLN 30-JUL-2003
 LOCUS AF284065
 DEFINITION Sesumum indicum myo-inositol 1-phosphate synthase mRNA, complete cds.

```

ACCESSION AF284065
VERSION AF284065.1 GI:3658815
KEYWORDS
SOURCE
ORGANISM Sesumum indicum (sesame)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterids; Lamiales; Pedaliaceae; Sesamum.
REFERENCE
AUTHORS Chun,J.A., Jin,U.H., Lee,J.W., Yi,Y.B., Hyung,N.I., Kang,M.H.,
Pyee,J.H., Suh,M.C., Kang,C.W., Seo,H.Y., Lee,S.W. and Chung,C.H.
Isolation and characterization of a myo-inositol 1-phosphate
synthase cDNA from developing sesame (Sesumum indicum L.) seeds:
functional and differential expression, and salt-induced
transcription during germination
JOURNAL Planta 216 (5), 874-880 (2003)
MEDLINE 22511750
PUBMED 12624775
REFERENCE 2 (bases 1 to 1845)
AUTHORS Jin,U.-H. and Chung,C.-H.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2000) Division of Biotechnology, Faculty of Life
Science and Resources, Dong-A University, Ha-Dan-Dong, Sa-Gu-Gu,,
Pusan City, Pusan 604-714, South Korea
FEATURES
source
1..1845
/organism="Sesumum indicum"
/mol_type="mRNA"
/db_xref="taxon:4182"
/feature_type="Seeds"
/feature="authorily: Sesumum indicum L."
51..1593
/codon_start=1
/product="myo-inositol 1-phosphate synthase"
/protein_id="AAG01148.1"
/db_xref="GI:3658816"
/translation="MFISFVESPVNYKTEGSHSVYVNETTLEHESRNGTYOMIV
KPKVYKFEKTDHVPKGLVNLVWGNNNGSTLTCGVANENEGISMAKRDVQDNF
GSLVKSIRVSPFNGEIVAPFKSLPVPNDVDFPGMDISNNALDAMGRAYLD
IDLOKLABPMEHVPPLPGIYDPDPIANOSRANVYKGRKEQVOOIIKMDRPEKE
ONKDKRVYVLTANTERYSNVYVGLNPDSELSMSVENEIETSTYVAYACVENY
PFNKSPONTVPGIIDLAIQNSLIGDDPDKSOTKRSYLVDPLVAGIKPSTIVS
YNHLGNNDGNNLSAPQTFRSKEIISKSNVVDVWVNSGILYBGEHPDHI VAKVYVY
GDSKRAMDEYSEIPIFMGKSTIVLHNTCEDSLAPILIDVLALNELSTRQLKAGE
GKFHSFHVAVATILSYLTRAPVYPGT PVVNALSKRAMLENILRACVGLAENNMILE
YK"
ORIGIN
Query Match 60.5%; Score 1064.2; DB 8; Length 1845;
Best Local Similarity 80.7%; Pred. No. 6.4e-276;
Matches 1243; Conservative 0; Mismatches 298; Indels 0; Gaps 0;
Qy 47 AAAAAATATGTTCAATCGAGAAATTTTAAGTTAAGTGTCTTAATGTGAAATACACGAGAC 106
Db 44 AAGCAAAATGTTCAATCGAGAGCTTCAAGGTTGAGAGCCCAATGTCAAGTTACACGAGGG 103
Qy 107 TGAGATTCAAGCGGTGTAACAATACGAAACCAACCAACTGTTTTCAGAGAACGAGATG 166
Db 104 TGAGATTCACTGTCTACAACTACGAGACTACAGAGCTTGTTCATGAGAGCAAGAAAGG 163
Qy 167 CACTATCAGTGTGTTTCAAAACCAAAATCTGTCAAAATACGAATTTAAAAACCAATCA 226
Db 164 CACTATCAGTGTGTTTCAAAACCAAAATCTGTCAAAATACGAATTTAAAAACCAATCA 223
Qy 227 TGTTCCTAAATTAAGGGGTATGCTTGTGGGTGGGGTGAACAACCGCTCAACCTCAG 286
Db 224 TGTTCCTAAATTAAGGGGTATGCTTGTGGGTGGGGTGAACAACCGCTCAACCTCAG 283
Qy 287 CGTGTGTTTATGCTTAACCGAGAGGCAATTCATGGGCTACAAAGACCAAGATTCACA 346
Db 284 TGGCGGTGCATTTGCAAAAGGGAAGGATTTCAAGGGCAACAAAGATTAAGGTGACACA 343
Qy 347 AGCCAATTAATTTGGCTCCTCCTCACCAAGCCTCAGCTATCCAGATTGGGTCTTCCAGGG 406

```

Db 344 GGCATTTATTTGGGCTGACGACAGGATCTCTCAATCAGAGTTGTTCTTCAATGG 403
 Qy 407 AGAGAAATCTATCCCATTCAGAGCCTGCTTCCATGTTAACTTGACGACATTTGT 466
 Db 404 AGAAGAGATCTATGCCCCCTTCAAAAGCCCTTCTCCATGTTCAACCCAGATAGT 463
 Qy 467 GTTTGGGGGATGGGATATCGCAACATGAACTGGCTGATGCCATGGCCGAGGAAAGT 526
 Db 464 GTTTGGGGGATGGGACATCGCAATGAAATTTGGCTGATGCCATGGGCGCAAGT 523
 Qy 527 GTTTGACATGATTTGACAGAGAGGATGAGCCTTACATGAAATCCATCTCCACTCC 586
 Db 524 GTTAGATATTTGATCTCCAGAGAGAGCTCAGGCCCTTACATGAGACATGTTCCCACTCC 583
 Qy 587 CGGAATCTATGACCCCGGATTTCAATGCTGCGCAACAGAGAGCGTCCCAACGTCAT 646
 Db 584 TGGAACTCTAGATCTGATTTCAATGAGCCGCAATCAGGGATTCAGTCCCAACGTCAT 643
 Qy 647 CAAGGCAACAAAGCAAGAGCAAGTTCACAAATCATCAAAAGATCAAGCGTTTAAAGA 706
 Db 644 TAAAGAACCCAAAGAAAGAACAGTTCAACAGATCATCAAGATGAGGAGCTTTAAGA 703
 Qy 707 AGCCACCAAGTSGACAAGGTGGTGTGTACTGTGCACTGCGCAACAGAGAGTACAGTAA 766
 Db 704 GCAAAATTAAGTSGACAAGGTGCTAGTGTGTGCAAGCCACAGAGAAAGATACAGTAA 763
 Qy 767 TTTGTTGTGGGCTTAAATGACAACATGAGAAATCTTTGGCTGTGTGACAGAAATGA 826
 Db 764 TTTGTTGTGGGCTTAAATGACAACATGAGAAATCTTTGGCTGTGTGACAGAAATGA 823
 Qy 827 GGTGATGATTTCTCTTCCATCTGTATGTCATGCTGTGTGTTATGAAATGTTCTT 886
 Db 824 GGGCGAATATCTCTTCAACCTTGAATGCAATGCTGTGTGTTGAAATGTTCTT 883
 Qy 887 CATTATGGAAGCCCTCAGAACATTTTTCACAGGGCTATGATCTTCCATCCGAG 946
 Db 884 CATTATGGAAGCCCAACAAATCTTTTTCACAGGGCTATGATCTTCCATCCGAG 943
 Qy 947 GAACACTTTGATTTGTGAGATGATCTTCAAGATGCTCAGACCAAAATGAAATCTGT 1006
 Db 944 GAACACTTTGATTTGTGAGATGATCTTCAAGATGCTCAGACCAAAATGAAATCTGT 1003
 Qy 1007 GGTGATTTCTTCTTGGGGCTGTATCAAGCCAACTATATGTCATTAACCATCT 1066
 Db 1004 AGTTGATTTCTTCTTGGAGCTGTATCAAGCCAACTATGTCATTAATCATCT 1063
 Qy 1067 GGGAAACAATGATGATGAAATCTTTCGGCTCACAACATTTCCGTTCCAAAGAAATCT 1126
 Db 1064 GGGAAACAACGAGGATGATTTATCAGCAACCCCAACTTTCCGTTCCAAAGAAATCT 1123
 Qy 1127 CAAGACCAACGTTGTGATGATATGTCACACAGCAATGCCATCTCTATGAGCTGTGA 1186
 Db 1124 TAAAGCAATGTTGTGAGACGATGTTGTCACCAACGCGATCTTTATGAGCTGTGA 1183
 Qy 1187 ACATCCAGACATGTTGTTTATTAAGTATGTCCTTACATGAGGAGACGACAGAGAG 1246
 Db 1184 ACACCCGACCAATGTTGTCTATCAAGTATGTAACGTAATGAGGAGGAGACGACAGAGAG 1243
 Qy 1247 CATGATGATGATCACTTCAGAGATTTCAATGGGTGAGAAAGAGACATGTTTGCACAA 1306
 Db 1244 AATGAGAGATACATTCGAGATATTCATGAGAGGAGGAGAAAGTACATGTTTGCACAA 1303
 Qy 1307 CACATCGAGATTTCCCTTTAGCTGCTCTATATCTTGGACCTTCTCTTCTGTA 1366
 Db 1304 CAGGTGAGAGATCTCTTCTGCTGCTGCTCATCTCTGAGATTTGCTCTCTAGCTGA 1363
 Qy 1367 GCTCAGACATGAAATGAGTTTAAAGCTGAAAATGAGGAAAATTCATCTCATTCACCC 1426
 Db 1364 ACTCAGACATCTGATTCAGCTCAAAAGCCGAGGAGAGGAGGCAAAATTCATCTCTCAACC 1423
 Qy 1427 AGTTGTACATCTGAGTACTCTCAAGAGGCTCTCTGTTCCACCGGATCACAGAT 1486

Db 1424 GGTGCAACATCTCAGTACTTAACCAAAAGCCCTTTGATCTCAGAGCACCGGT 1483
 Qy 1487 GGTGAATGCATTTCTAAAGCAGCGTGCATGCTGGGAAAATATATAGGGCTTGTGG 1546
 Db 1484 CGTCAACGACATCTCTAACAGCGTGCATGCTCGAACAATCTTGAAGGCTTGTGG 1543
 Qy 1547 ATTGCCCGCAGAGAAATACATGATTTCTCGAGTACAGTGA 1587
 Db 1544 ATTGCTCCAGAAAACATATGATTTTGAATATCAAGTGA 1584
 RESULT 12
 MCU32511
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mesembryanthemum crystallinum (common iceplant)
 Mesembryanthemum crystallinum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Caryophyllales; Astaceae; Mesembryanthemum.
 1 (bases 1 to 2053)
 Ishitani, M., Majumder, A. L., Bornhouser, A., Michalowski, C. B.,
 Jensen, R. G., and Bohnert, H. J.
 Coordinate transcriptional induction of myo-inositol metabolism
 during environmental stress
 Plant J. 9 (4), 537-548 (1996)
 96208959
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 8624516
 2 (bases 1 to 2053)
 Ishitani, M., Majumder, A. L., Bornhouser, A., Michalowski, C. B.,
 Jensen, R. G., and Bohnert, H. J.
 Direct Substitution
 Submitted (27-JUN-1995) Hans J. Bohnert, Biochemistry, University
 of Arizona, Biosciences West 516, Tucson, AZ 85721, USA
 FEATURES
 Location/Qualifiers
 1..2053
 /organism="Mesembryanthemum crystallinum"
 /mol_type="mRNA"
 /db_xref="taxon:3544"
 182..1720
 /codon_start=1
 /product="myo-inositol-1-phosphate synthase"
 /protein_id="AA03687.1"
 /db_xref="GI:975888"
 /translation="MTESFKYSPVRYKTEIEISVYVYDTTELVHNRKAGYOW
 IYKPTVQYHFRITDRPRLGVMVGWGNSSITLGGVIARBEISATKDKIQOAN
 YFGSLTQASISIVGSPNGEIVAPKSLIPWNPDPVVGWDIDSMNLADMTARAV
 PDIDQKQLRPVMEHVPLPGIYDDPIAANGSANNVYIKETKEOVERYKDIREF
 KEKNKYDVLVMTGNTERYSNVYVDDIENMLASLEKSEISPSVALACIES
 NIPITNGSPQMFVPEGLDIAIKNSLIGDGFSGQOTKRSVLVDFVAGIKPISI
 VSNYHGNDDGNLSAPQTFRSEIKSVYVDWMASSNIIYEPREHPIVYIKVP
 YVGRKRAVDERTSIFMGNTIVWANTCSLSLAAPILDLVLAELSTRIQLAE
 EEDKPSFHPVATILSYLTKAPLVPPGPVVAALSKORAMLENILRACVGLAPENMI
 LEYK"
 ORIGIN
 Query Match 59.0%; Score 1037.8; DB 8; Length 2053;
 Best Local Similarity 77.6%; Pred. No. 8.7e-269;
 Matches 1282; Conservative 0; Mismatches 362; Indels 7; Gaps 2;
 Qy 44 GTGAAAATAATGTCATGAGAAATTTTAAAGTTGAGTGTCTCTAAATGTAAGTACACGA 103
 Db 172 GAGAGAGAAAATGTTTATGAGAGCTTCAAAGTAGAGAGCCAAAGCTGAATGACAGA 231
 Qy 104 GACTGATTCAGTCTGCTGATCACTACGAAACCAACCACTGTTTACAGAGAACAGA 163
 Db 232 GAATGATTTGATCGGTGTACATGACACCAACCGAGTTGGTTATCATGAGATGCAA 291
 Qy 164 -----TGGCACTATCAGTGATGTCAAAACCAATCTGTCAAAATGCAATTTAAAC 217

D	b		292	AGATGCTGGTGTTTATCAATGAGTCGTAAAGCCCAAGACCGTTTCAGTAATCATCTTCAAGAC	351
O	y		218	CAACATCCATGTTCTCTAAATTGAGGGGTATATGCTTGTGGGTGGGGAACAACGCGCTC	277
D	b		352	TGACACTCGTGTCGCCAAGACTCGGGGTGATGCTGTTGGATTGGGGTGGAAATAACGCGCTC	411
O	y		278	AACCTTCACCGGTGTGTATTGTCTAACCGAGAGGCATTTTCATGGGCTTCMAAGGACA	337
D	b		412	CACCTTCACAGGTGTGTGATCGCCAACCGGAGGGAAATCTTTGGCCAACCAAGGACAA	471
O	y		338	GATTCAAAGAACCATACTTGTGGCTCCCTCACCCAGACCTCAGCTATCCGAGTGGGTC	397
D	b		472	GATCCAGACAGCACTATTTTTTGGTTGTTGACCCAGGCTTTCTTAATTGGGTCGGGTC	531
O	y		398	CTTCCAGGAGAGGAAAICTATGCCCCATTCAGAGCGCTGTTCCAAATGGTTAACCTTA	457
D	b		532	TTTAAATGGAAGAGAAATTATGCTCCCTTCAAGAGCTTGCTCCCTATGGTAAACCGA	591
O	y		458	CGACATGTGTTTGGGGATGGGATATCAGCAATGAACTTGGCTGATGTCATGGCCAG	517
D	b		592	TGACGTAGTGTTTGGGGTTGGGACATPAAGGCAGATGAACCTGTGCTATGTAATGAACAG	651
O	y		518	GGCAAGGTGTTTACATTCGATTTTGCAAGACAGTGAAGGCTTACATGGAATCATGCT	577
D	b		652	GGCTAGGGTCTTTTACATTTGATCTTCAAAACAGACTCAGGCTTACATGGAGCATGCT	711
O	y		578	TCCATCCCGCCGAATCTATGACCCCGAATTTTCATGTGCTGCNAACCAAGAGAGCGTGCAA	637
D	b		712	TCCCTCCCTGGTATTATACGACCGTGAATTCATGCTGCGAAGCAAGGCTCCGAGCGAA	771
O	y		638	CAACGTCAATCAAGGCGACAAAGCAAGCAAGTTCAACAAATCATCAAAGACATCAAGGC	697
D	b		772	CAACGTCAATTAAGGGCACCAAGAAAGGAACAAGTTAGAGAGTCAATTAAGCATTTAAGGA	831
O	y		698	GTTTAAAGGAACCCACCAAGTGGACAAGAGTGTGTACTGTGGACTGCGAACACAGAGAG	757
D	b		832	GTTCAAGGAGAAACAAGTGGACAAAGTGTGTGTATGACTGGACCTGGACCAACAGAGAG	891
O	y		758	GTACAGTAAATTTGGTTGGGCTTTAATGACACATGAGAGATCTTTGGCTGCTGTGA	817
D	b		892	GTACAGCATGTGTGTGTGCGGGCTCAACGACACATGAGAAACCTGTTGGCATCTTGGGA	951
O	y		818	CAGAAATGAGCTGAGATTTCTCCTTCACCTTGATGCCATGCTGTGTGTAAGAAAA	877
D	b		952	GAAAAACAGTCGAGGATTTCTCCATTCACATCTCGATGCTTGGCAGTAGTATCGAAGAAA	1011
O	y		878	TGTTCCCTTTCATTAATGGAACCCCTCAAGACATTTGTATCAGGGCTGATGTATCTTGC	937
D	b		1012	TATTCCTTTCATCAATGGTAGCCCTCAACAAACCTTTGTGCAAGGGCTCATTTGATGTGGC	1071
O	y		938	CATCGCGAGAAACCTTGTGATTTGGTGGAGATGACTTCAAGAGTGGTCAGACCAAAATGAA	997
D	b		1072	TATCAAGAAAGAACTTTGATTTGGTGTGTGATGACTTCAAGAGTGGTCAGACCAAGATGAA	1131
O	y		998	ATCTGTGTTGTTGATTTCTTGTGGGGGCTGTATCAAGCCAAATCTATATGTCAAGTTA	1057
D	b		1132	ATCTGTCTCGTCGACTCTCTGTGTGGGGCTGTATTAAGCCAACATCATATGTAAGCTTA	1191
O	y		1058	CAACCATCTGGGAAACATGATGTGATGTAATCTTTCCGCTCCACAAATTTCCGTTCCAA	1117
D	b		1192	CAACCATCTGGGAAACATGATGGGATTAACCTGTCAAGCACACTTCAAACTTTCAGGTCCA	1251
O	y		1118	GGAAATCTCAAGAGCAAGTTGTGATGATATGGTTCAAAGCAATGCGATCTCTATGA	1177
D	b		1252	GGAGATTTCTTAAGACAACTGTGTGATGACATGTTGTCCAGCAACGGCAATCTCTATGA	1311
O	y		1178	GCTGTGTAAACATCCAGACCAATGTTGTGTTATTAATGATGTGCTTATCGTAAGGGGACAG	1237
D	b		1312	GCGGGGCGAAACCTGTATCATGTTGTGTGTATCAAGATATGTGCAATATGTGGGAGACAG	1371
O	y		1238	CANAGAGCCATGATGATGATCACTTCAAGATATTTCAATGGGTGAAAGACACATTTGT	1297

[illegible]

ORIGIN

DSKRANDEYSEIFMGKSTIVMHTCEDSLAAPIILDLVILLAEISTRITOLKAESEG
KPHSFHVAITILSYLTKAPLVPGTPVVALNSKORAMENIIRACVGLAPENNMLEY
K"

Query Match 58.9%; Score 1037.4; DB 8; Length 1845;
Best Local Similarity 80.0%; Pred. No. 1.1e-268;
Matches 1234; Conservative 0; Mismatches 306; Indels 3; Gaps 1;

QY 44 GTGAAAAAATATGTTTANTGAGAAATTTTAAGTTGAGTGTCTTAATGTGAAGTACCGA 103
DB 17 GAGAGAAAATATGTTTATGAGAGCTTCAAGGTTGACAGCCCAAGCTTAAGTACAG 76
QY 104 GACTGATTCCAGTCCGTGACAACTACGAAACACCGAACTTGTCAGAGAAACGAA 163
DB 77 CAGTGAGATTCTCCGTCTACACATATGAAACTGCAAGCTTGTCAGAGAAACGAA 136
QY 164 TGGACCTTACAGTGAATGTCAAACCAATCTGCAATACGAATTTAAACCAACAT 223
DB 137 CGGTACTTACAGTGAATGTCAAACCAATCTGCAATACGAATTTAAACCAACAT 196
QY 224 CCATGTTCTTAATTTAGGGTAAATGCTTGGGTTGGGTTGGGTTGGGTTGGGTTGGG 283
DB 197 CCATGTTCTTAATTTAGGGTAAATGCTTGGGTTGGGTTGGGTTGGGTTGGGTTGGG 256
QY 284 CACCGGTGGTGTATTTGCTTAAACGAGAGGCAATTTGATGGGCTTCAAGAGCAAGTTGA 343
DB 257 GACAGGTGGTGTATTTGCTTAAACGAGAGGCAATTTGATGGGCTTCAAGAGCAAGTTGA 316
QY 344 ACAAGCCAAATTAATTTGCTTCAACCAAGCCGATCCGATCCGATCCGATCCGATCCGAT 403
DB 317 ACAGGCCAAATTAATTTGCTTCAACCAAGCCGATCCGATCCGATCCGATCCGATCCGAT 376
QY 404 GGGAGAGAAATCTATGCCCCATTCAGAGCTGCTTCAATGTTAAACCTTGACACAT 463
DB 377 TGGAGAGAAATCTATGCCCCATTCAGAGCTGCTTCAATGTTAAACCTTGACACAT 436
QY 464 TGGTGTGGGGGATGGGATATACGAAACATGAACTGGGTTGATGGGTTGATGGGTTGAT 523
DB 437 AATTTTGGGGGATGGGATATACGAAACATGAACTGGGTTGATGGGTTGATGGGTTGAT 496
QY 524 GGTGTTTGAATGATTTGTCAGAGAGAGTGAAGGCTTATCATGGAATTCATGCTTCCACT 583
DB 497 AAGGTGATGATGATTTGTCAGAGAGAGTGAAGGCTTATCATGGAATTCATGCTTCCACT 556
QY 584 CCCCCGAAATCTATGAGCCGGAATTCATGCTGCAACCAAGAGAGCGTCCCAACAGCT 643
DB 557 CCTGGGATCTACTACCTGACTTCACTGCTGCAATCAAGGCTTCAAGGCTTCAAGGCTT 616
QY 644 CATCAAGGCAACCAAGAGAGAGTGAAGGCTTATCATGGAATTCATGCTTCCACT 703
DB 617 GATCAAGGCAACCAAGAGAGAGTGAAGGCTTATCATGGAATTCATGCTTCCACT 676
QY 704 GGAAGCCAAATGAGAGAGTGAAGGCTTATCATGGAATTCATGCTTCCACT 763
DB 677 GGAAGAAATTAAGTGAAGAGTGAAGGCTTATCATGGAATTCATGCTTCCACT 733
QY 764 TAAATTTGTTGGGCTTAAATGACACATGGAATTCATGCTTCCACT 823
DB 734 CAATGTTGTTGGGCTTAAATGACACATGGAATTCATGCTTCCACT 793
QY 824 TGAAGGTGATTTTCCCTTCAACCTGATATGCAATGCTTGTGTATGGAATTTGCTTCC 883
DB 794 TGAAGGTGATTTTCCCTTCAACCTGATATGCAATGCTTGTGTATGGAATTTGCTTCC 853
QY 884 TTTCAATTAATGAGAGCCCTCAGAACATTTTATACAGAGGCTGATGATCTGCGCATCGC 943
DB 854 TTTTATCATGAGAGCCCAACCAACCTTGTGTCAGAGGCTGATGATGATGATGATGATGAT 913
QY 944 GAGGAACCTTTGATTTGAGAGATGATCTTCAAGAGTGTGACCAACCAATGAAATCTGT 1003
DB 914 AAGGAACCTTTGATTTGAGAGATGATCTTCAAGAGTGTGACCAACCAATGAAATCTGT 973

QY 1004 GTTGTGATTTCTTCTTGGGGGCTGTATCAAGCAACATCTATAGTACAGCA 1063
DB 974 GCTAGTTGATTTCTTCTTGGGGGCTGTATCAAGCAACATCTATAGTACAGCA 1033
QY 1064 TCTGGGAAAATATGATGATGATGATTTTGGGCTTCAACAACTTTCCGTTCAAGGAAAT 1123
DB 1034 TTTGGGAAAATATGATGATGATGATTTTGGGCTTCAACAACTTTCCGTTCAAGGAAAT 1093
QY 1124 CTCCAAGAGCAAGTGTGATGATGATTTTGGGCTTCAACAACTTTCCGTTCAAGGAAAT 1183
DB 1094 CTCCAAGAGCAAGTGTGATGATGATTTTGGGCTTCAACAACTTTCCGTTCAAGGAAAT 1153
QY 1184 TGAACATCCAGACATGTTGTTGTTATTAAGTATGATGATGATGATGATGATGATGAT 1243
DB 1154 AGAGCACTTGATCATGTTGTTGTTATTAAGTATGATGATGATGATGATGATGATGAT 1213
QY 1244 AGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1303
DB 1214 AGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1273
QY 1304 CAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1363
DB 1274 CAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1333
QY 1364 TGAAGTCAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1423
DB 1334 GGAATCAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1393
QY 1424 CCCAGTGTACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1483
DB 1394 TCCCGTGGCAGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1453
QY 1484 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1543
DB 1454 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1513
QY 1544 TGAATTTGGGCGCAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1586
DB 1514 TGAATTTGGGCGCAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1556

RESULT 14
AF433879
LOCUS 1986 bp mRNA linear PLN 07-NOV-2001
DEFINITION Sueda maritima subsp. salina myo-inositol-1-phosphate synthase
(INPS) mRNA, complete cds.
ACCESSION AF433879
VERSION
KEYWORDS
SOURCE Sueda salina
ORGANISM Sueda salina
REFERENCE Wang, J.P., Wang, P.P., Sun, Y.F., Zhao, Y.X. and Zhang, H.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2001) The Biology Department of Shandong Normal University, Key Laboratory of Plant Stress Research, No.88, Wenhua East Road, Jinan, Shandong Province 250014, China

FEATURES
source location/Qualifiers
1..1986
/organism="Sueda salina"
/mol_type="mRNA"
/sub_species="salina"
/db_xref="taxon:126914"
1..1986
/gene="INPS"
140..1672
/gene="INPS"
/codon_start=1
/product="myo-inositol-1-phosphate synthase"
/protein_id="AAU28131.1"

```

/db xref="GI:16755889"
/translation="MEIENFKVESPNVNTENEIHSVYVYETTELVHENRKGSGYOMI
VPEKTVKQFOKTOHPKLVGMVJVGSGNCGCTLGGVYANEGSGSMATKXIOOANY
RSLTOASSIRVSGNGBEITVAPEKSLPMVNDPIVRGKMIISLNLADMTARAKVL
DIDLOKOLRPVMEHWPIBGIYDPPFIANQSRANNIIKTKRQOEVOYIDJHEPK
ENKVDKVVLTWNTANTERYSDVVGJLNTDENLMAVAENESIEPSTLYATACIOED
VPIFGSPONTFPIGLIELAIKNCLEIGDDPKSQVQLVPLVFGAKRPSIV
SYNHLGNDGNMLSAPOFRSKIEISKNVDDMVASNGILYEPGEHPDHVVYKVPY
VDSKRAMDEYSEIEMGSKNTIVLNTCESLAAPIILDVLAEISTRLEFAET
BDKPHFHFHVATILISYLTKAPLVPPTPVNALSQRMLBQRACVGLAPBNMILE
YK"

```

ORIGIN

```

Query Match      58.5%; Score 1030; DB 8; Length 1986;
Best Local Similarity 80.2%; Pred. No. 1.1e-266;
Matches 1236; Conservative 0; Mismatches 300; Indels 6; Gaps 2;

```

```

QY 50 AATATGTCATCGAGAAATTTTAAGTTGAGTGTCTTAATGTGAGTACACGAGACTGA 109
DB |||||
DB 136 AAAAAATGTCATCGAAAAATTTCAAGGTAGAAAAGCCCAATGTTAACTACAGTGAATGA 195
QY 110 GATTCAGTCGCTGTCACATGAGAAACCAACCAACTGTTCACGAGAACAGGAATGGAC 169
DB |||||
DB 196 AATTCATTAAGTTACAAATTATGAAACCACTGAATTAATTCACGAGAAATCGCAAAGCTC 255
QY 170 ---CTATCAGTGGATGTGCAAAACCAATCTGTCAATAGCAATTTAAACCAACATCCA 226
DB |||||
DB 227 TGTTCCTAAATTAAGGGGTAATGCTGTGGGTGGGTGGAACCAACGCTCAACCTCAC 286
QY 316 TGTTCCTAAATTAAGGGGTAATGCTGTGGGTGGGTGGAACCAACGCTCAACCTCAC 375
DB |||||
DB 287 CGTGGTGTATTTGCTAACGAGAGGSCATTCATGSGGCTAACAGAAAGAAAGTTCAAA 346
QY 376 CGTGGTGTATTCGCAATCGCAGAGGTAATATATGAGCAACAAAGAAACAAATTAACA 435
DB |||||
QY 347 AGCAATTAATTTGGCTCCCTCAACCAACCTCAGTATCCGAGTTGGGTCTTCACAGG 406
DB |||||
DB 436 AGCTAATCTATTTGAGATTCCTCACTCAGGCTTCTTATTCGTGTGGTCTTTTAATGG 495
QY 407 AAGAGAAATCTATGCCCCATTCAAGAGCCTGCTTCAATAGTTAACCTCGACGACATTTGT 466
DB |||||
DB 496 AAGAGAGATTTATGCTCTTTTAAAGACCTTCTTCTTATGAGTGAACCCAGATGATATAGT 555
QY 467 GTTTGGGGATGGGAATATCAGCAACATGAACCTGGCTGATGCCATGGCCAGGCAAGGT 526
DB |||||
DB 556 GTTTGGGGGCTGGGACATAGTGAACCTGAGACAGCCATGACGAGGGCAAAAGGT 615
QY 527 GTTTGACATCGATTTTGCAGAGAGAGTGAAGCCTTACATGGAATCCATGCTTCCACTCCC 586
DB |||||
DB 616 GTTTGACATCGATTTTGCAGAGAGAGTGAAGCCTTACATGAGACATAGTTCCTCCCTTCC 675
QY 587 CGGAATCTATGACCCGGAATTTCTTGTCTGCAACCAAGAGAGCGTTCACCAACGTCAT 646
DB |||||
DB 676 TGGTATCTATGACCCGGAATTTCTTGTCTGCAACCAAGATTTCTGCGAAATTAACATCAT 735
QY 647 CAAGGCAACAAGAGAGAGAACTTCAACAAATCAACAAAGACATGAAGCCGTTTAAGA 706
DB |||||
DB 736 TAAAGGACACTTAAGAGAGAGCAATGAGCAAGCATTAAGAGACATGAGGAATTTAAGA 795
QY 707 AGGCACCAAAAGTGAACAAGGTGTTGATCTGTGAGCTGCAACACAGAGAGTACAGTAA 766
DB |||||
DB 796 GAAGAACAAGTGAACAAGGTGTTGATCTGTGAGCAAGCAACACAGAGAGTACAGTGA 855
QY 767 TTTGTTGTGGGCTTTAATGACCAATGAGAAATCTTTGGCTGCTGTGAGACGAATGA 826
DB |||||
DB 856 TGTGTTGTGGGCTCAATGACCAACAGAGAACTTATGAGCAGCTGTTGAAGAATGA 915
QY 827 GGTGAGATTTCTCTTCCACTTGTATGACATTTGCTGTGTTATGAAAAATGTTCTTT 886
DB |||||
DB 916 GTCTGAGATTTCTCAATGACATTTGATGATGCTTGCATCAAGAGAGATGCTCTTT 975

```

```

QY 887 CATTAATGAAGCCCTCAAGAACCTTTTGATCAAGGGCTGATGATCTTGGCATGCGGAG 946
DB CATCAACCGTATGCTCTCAAGAACCTTTTGTCTGCTGCTCATTTGAACCTGGCCATAAAGA 1035
QY 947 GAACCTTTGATGGTGAAGATGATCTTCAAGAGTGGTCAGACCAAAATGAATCGTGT 1006
DB |||||
DB 1036 GAATGTTGATGGAGAGATGATCTTCAAGAGTGGTCAGACCAAAATGAATCGTGT 1095
QY 1007 GGTGATTTCTTGTGGGGGCTGTATCAAGCCCAATCTATAGTCAGTTACAACTAT 1066
DB |||||
DB 1096 CGTTACCTTCTTGTGGGGGCTGTATTAAGCAACATCAATTTGAGCTCAACACACTT 1155
QY 1067 GGGAAACATGATGATGATGATCTTGGGCTTCACAACTTTCCGTTCCAGAGAAATCTC 1126
DB |||||
DB 1156 GGGAAACATGATGATGATGATGATCTTGGGCTTCACAACTTTCCGTTCCAGAGAAATCTC 1215
QY 1127 CAAGGCAACGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186
DB |||||
DB 1216 AAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
QY 1187 ACATCCAGACCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1246
DB |||||
DB 1276 ACACCTGATCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1335
QY 1247 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306
DB |||||
DB 1336 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1395
QY 1307 CACATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366
DB |||||
DB 1396 CACCTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1455
QY 1367 GCTCAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1426
DB |||||
DB 1456 GCTTACGACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1515
QY 1427 AGTTCATACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1486
DB |||||
DB 1516 AGTTCATACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1575
QY 1487 GGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1546
DB |||||
DB 1576 GGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1632
QY 1547 ATTGGCCCAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1588
DB |||||
DB 1633 TTTAGCTCCAGAGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1674

```

RESULT 15
AY065415
LOCUS
DEFINITION
Arabidopsis thaliana putative myo-inositol-1-phosphate synthase
(At5g10170) mRNA, complete cds.
ACCESSION
AY065415
VERSION
AY065415.1 GI:17529257
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1863)
REFERENCE
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banb, J., Chung, M.K.,
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,
Carinci, P., Chen, H., Cheuk, R., Hayashitaki, Y., Ishide, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K.,
Davis, R.W., Ecker, J.R. and Theologis, A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 1863)

AUTHORS	Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, P., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
TITLE	Submitted (04-DEC-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
JOURNAL	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
COMMENT	The Salk, Stanford, PGENC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.
FEATURES	Yamada, K. (SSP/PGENC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGENC) contributed equally to this work as PIs.
source	Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.
Location/Qualifiers	1. .1863 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="5" /clone="RAF109-36-B07 (R18590)" /ecotype="Columbia" /note="This clone is in a modified pBluescript vector (pUC-1) as a BamHI/XhoI insert." 1. .1863 /gene="At5g10170" 1. .90 /gene="At5g10170" 91. .1623 /gene="At5g10170" /codon_start=1 /evidence=experimental /product="putative myo-inositol-1-phosphate synthase" /protein_id="AAL38856.1" /db_xref="GI:17529258" /rna_label="M1ESFVSPNPKYKTENEHSVVDYOTTELHKNKGAPOWTVKRTVYERKTDHVKLGLVLMVGMGNSSTLTAGVIANRGIISMAETKIOQANFSGLTQASSIRVGSFNGEETIYAPFKSLIPVNPPEEIVFGSDISDNMLAAMAKALDIDLQKMRPFMEHWPLPGIFDPDTIAAQSSNANVIGTKKQOLEVYIKOIRKEKRRKVDVAVLWANTERISNVVGLNDITENLMSLDKDEAEISPELTIAICVLEVPFINSPOMTFVGLIELAIKRNCLIGDDFKSGTKMSVLDVFLVAGIKPTSVSNHLGNNDMLNAPOTPRSKISKNVVDVGSNGILYBGEHPDHVAVKYPCVGSKRAMDEYTSIEIPGKQKTIYMHNTGDSILAAPILIIDIVLIELTRIPQMSNEGRFHSFHPVATLISYLSKAPLVPFGTIPVVALSKORAMLENVLRACVGLAPENNMILEYK"
gene	1. .1863 /gene="At5g10170"
5' UTR	1. .90 /gene="At5g10170"
CDS	91. .1623 /gene="At5g10170" /codon_start=1 /evidence=experimental /product="putative myo-inositol-1-phosphate synthase" /protein_id="AAL38856.1" /db_xref="GI:17529258" /rna_label="M1ESFVSPNPKYKTENEHSVVDYOTTELHKNKGAPOWTVKRTVYERKTDHVKLGLVLMVGMGNSSTLTAGVIANRGIISMAETKIOQANFSGLTQASSIRVGSFNGEETIYAPFKSLIPVNPPEEIVFGSDISDNMLAAMAKALDIDLQKMRPFMEHWPLPGIFDPDTIAAQSSNANVIGTKKQOLEVYIKOIRKEKRRKVDVAVLWANTERISNVVGLNDITENLMSLDKDEAEISPELTIAICVLEVPFINSPOMTFVGLIELAIKRNCLIGDDFKSGTKMSVLDVFLVAGIKPTSVSNHLGNNDMLNAPOTPRSKISKNVVDVGSNGILYBGEHPDHVAVKYPCVGSKRAMDEYTSIEIPGKQKTIYMHNTGDSILAAPILIIDIVLIELTRIPQMSNEGRFHSFHPVATLISYLSKAPLVPFGTIPVVALSKORAMLENVLRACVGLAPENNMILEYK"
3' UTR	1624. .1863 /gene="At5g10170"
misc_difference	1846 /gene="At5g10170" /note="compared to genomic sequence" /replacement="c"
ORIGIN	
Query Match	57.4%; Score 1010.8; DB 8; Length 1863;
Best Local Similarity	78.3%; Pred. No. 1.7e-261;

Matches 1213; Conservative 0; Mismatches 337; Indels 0; Gaps 0;	
QY 50 AATAATGTCATGAGAAATTTTAAGTTGAGTCTCTAATGTGAAGTACCCGAGCTGA	109
DB 87 AAAAAAGTTCATGAAAGCTTCAAGGTGTGATCTCCGAACGTGAAGTACCCGAGATGA	146
QY 110 GATTCAGTCGCTGTACAACTAGCAAAACCCGAACTGTTTACAGAGAACGAAATGAC	149
DB 147 GATTCAGTCGCTGTACAACTAGCAAAACCCGAACTGTTTACAGAGAACAAACGCTGC	206
QY 170 CTATCAGTGGATTTGTCAAAACCCAAATCTGTCAATATGCAATTTAAACCAATCATGT	229
DB 207 TTTTCAATGAGACTGTGAAGCAAAAGCTGTGAAGTATGAGTTTAAACGAGCATATGT	266
QY 230 TCTTAATTAAGGGGTATGCTGTGTGGTGTGGGTGGAACACGGCTCAACCTCACCGG	289
DB 267 TCCCAAACTTGGAGTAACTCTGTTGTGGGAGGCAACAAATGTTTACATTTACCTCCG	326
QY 290 TGGTGTATTTGCTAACCGGAGGCGCATTTCAATGGCTTACAAAGGAGAAATTCAGAACG	349
DB 327 TGGTGTATTTGCTAACCGGAGGCGCATTTCAATGGCTTACAAAGGAGAAATTCAGAACG	386
QY 350 CAATTACTTGTGGCTCCCTCACCAAGCTCAGCTATCCGAGTTGGGCTTCCAGGAGAGA	409
DB 387 TAATTAATTTGGGTCTTAATCCCAAGCATCTTATTCGGGTGGGTCAATTAATGTGA	446
QY 410 GGAATCTATGCCCATTCAGAGCTGTCTCAATGTGTTAACTTGACGACATTTGTGT	469
DB 447 AGAATCTATGCTCTTCAAAAGCTTCTCCCATGTGGAATCCGAGAAATTTGTGT	506
QY 470 TGGGGATGAGATATAGCAATGAACCTGGCTATGCGATGCGCAAGGCAAAAGGTGT	529
DB 507 TGGAGATGAGCATTAAGTATGATGAATGATGATGATGATGATGATGATGATGATGAT	566
QY 530 TGACATGATTTGACAGAAAGTGTAGGCGCTTACATGATGATGATGATGATGATGATGAT	589
DB 567 AGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	626
QY 590 AATCTATGACCCGATTTGATGCTGCGCAACCAAGAGAGCGTCCCAACAGCTCATCA	649
DB 627 AATCTTGCAGCCAGATTTGATGCTGCGCAACCAAGAGAGCGTCCCAACAGCTCATCA	686
QY 650 GGGCAACAAAGCAAGCAAGTCAACAAATCATCAAGCAATCAAGCGCTTTAAAGAAAG	709
DB 687 AGGCAACAAAGCAACCAAGTCAACAAATCATCAAGCAATCAAGCGCTTTAAAGAAAG	746
QY 710 CACCAAGTGAACAAGGTGTGTACTGTGACCTGCGCAACAGAGAGGTACAGTAATT	769
DB 747 AACAAGTGAACAAGGTGTGTACTGTGACCTGCGCAACAGAGAGGTACAGTAATT	806
QY 770 GGTGTGGGCTTAAATGACACCAATGAGATCTTTGCTGTGTGACAGAAATGAGAG	829
DB 807 GGTGTGTGTCTTAAACGACACGAGAGATCTCATGTCTTTGACAGAGATGAGAG	866
QY 830 TGAATTTCTCTTCACTTGTATGACCTGCTGTGTATGAGAAATGTTCTTCTCAT	889
DB 867 TGAATTTCTCTTCTTCACTTGTATGACCTGCTGTGTATGAGAAATGTTCTTCTCAT	926
QY 890 TATGGAAGCCTCAAGCACTTTTGTACCAAGGCTGATGATCTTGGCATGCGCAGAA	949
DB 927 CATGGAAGTCAACAAACCTTTTGTCTTGGCTTATGATTAATGACATTAAGGAA	986
QY 950 CACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1009
DB 987 CTGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1046
QY 1010 TGAATTTCTTGTGGGCGCTGTATCAAGCAACATGTATGATGATGATGATGATGATGAT	1069
DB 1047 CGATTTCTTGTGTGGCGCGGATCAACCGACGCTGATGTGATGATGATGATGATGATGAT	1106
QY 1070 AACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1129
DB 1107 GAACAACGACGAGTGAATCTGTCAAGCACTTCAACATTTCCGCTCAAGAGATATTCGA	1166

```
QY 1130 GAGCAACGTTGTTGATGATATGATCAACAGCAATGCCATCTCTATAGACCTGTGAAACA 1189
    |||||
Db 1167 GAGCAATGTCGTGAGACGATATGTTGTGTAACAACGAAATTTGTAGAGCTGTGAAACA 1226
    |||||
QY 1190 TCCAGACCATGTTGTTGTTATTAAGTATGTCCTTAAGTAGGGGACAGCAAGAGCCAT 1249
    |||||
Db 1227 CCTGACCAAGTTGTGTGTCATTAAAGTATGTGCCATGTGTAGAGACAGCAAGAGCAAT 1286
    |||||
QY 1250 GGATGAGTACATTGAGAGATATTCATGGGTGAAAAGACCACTTTTGCACAACAC 1309
    |||||
Db 1287 GGATGAAATACACATCAGAGATATTCATGGGAGGGAACACAGATTGTGATGCACAACAC 1346
    |||||
QY 1310 ATGCGAGATTCCTCTTAGCTGCTGCTATTAATCTTGAAGTGTGCTTCTGTGAGCT 1369
    |||||
Db 1347 TTGTGAAGACTCTCTCTTGGCTGCTCAATTAATTTGAGCTTGTCTTCTTGTGAGCT 1406
    |||||
QY 1370 CAGCACTAGAAATCGAGTTTAAAGCTGAAAATGAGGAAAATTCACCTCATTCACCCAGT 1429
    |||||
Db 1407 CACTACTGGATCCAGTTCATGAGTGAACAGAGGAAAGTTTCATTTCCATCCGCT 1466
    |||||
QY 1430 TGGTACCATCTCAGCTACCTCACCAAGCTCTCTGTGTTCCACCGGTACACCAAGTGT 1489
    |||||
Db 1467 GGCTACTCTCTTAAGTTACCTCAAAAGGCACCACTGTTCTCTCGGACCGCCAGTTGT 1526
    |||||
QY 1490 GAATGATTTGTCAAGCAGCGTGCATGCTGGAACATATATGAGGGCTTGTGTGATT 1549
    |||||
Db 1527 TAATGCTTTGTGAAACAGCGTGCATGCTTGAAGACGTACTCAGAGCTTGTGTGACT 1586
    |||||
QY 1550 GGCCCCAGAGAAATTAACATGATTTCTCGAGTACAAAGTAGAGCATGGAGCCGA 1599
    |||||
Db 1587 TGCTCTGAGAACAAACATGATTTCTCGAGTACAAAGTAGATCAAGAGACCAA 1636
    |||||
```

Search completed: June 8, 2005, 02:53:56
Job time : 7692.08 secs